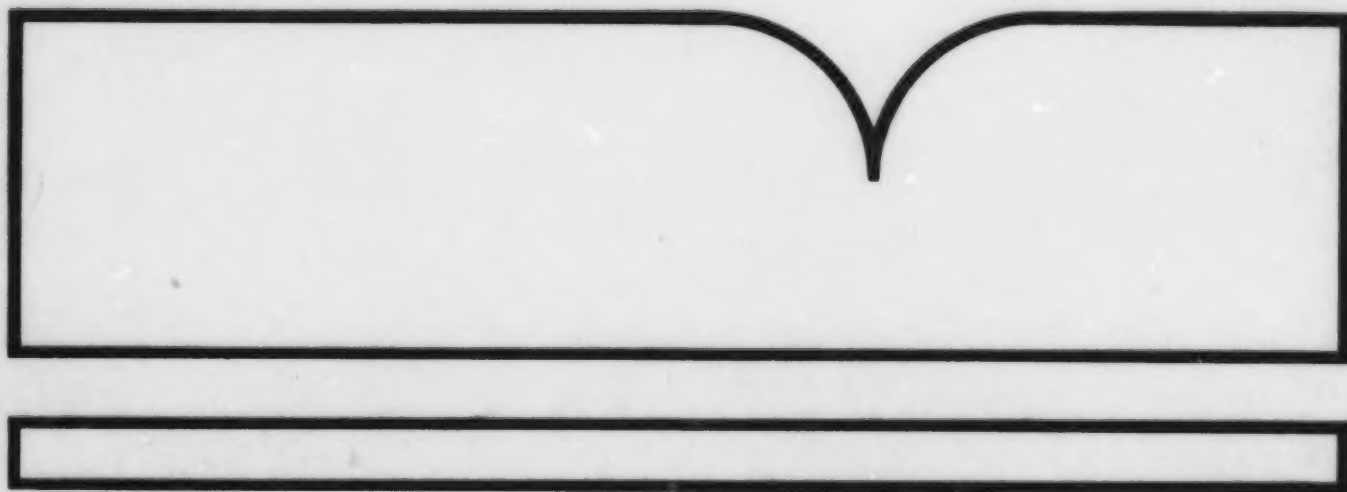


Evaluation of Remote Sensing Data for Input into
Hydrological Simulation Program-FORTRAN (HSPF)

Hydrocomp, Inc.
Mountain View, CA

Prepared for
Environmental Research Lab.
Athens, GA

Jun 81



U.S. Department of Commerce
National Technical Information Service

NTIS

TECHNICAL REPORT DATA (Please read instructions on the reverse before completing)		
1. REPORT NO. EPA-600/3-81-037	2. ORD Report	3. RECIPIENT'S ACCESSION NO. P001 209581
4. TITLE AND SUBTITLE Evaluation of Remote Sensing Data for Input into Hydrologic Simulation Program-FORTRAN (HSPF)	5. REPORT DATE June 1981	
	6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) D.D. Franz and S.M. Lieu	8. PERFORMING ORGANIZATION REPORT NO.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Hydrocomp, Inc. Mountain View CA 94040	10. PROGRAM ELEMENT NO. CARBLA	
	11. CONTRACT/GRANT NO. 68-01-5801	
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Research Laboratory--Athens GA Office of Research and Development U.S. Environmental Protection Agency Athens GA 30613	13. TYPE OF REPORT AND PERIOD COVERED Final, 3/79 - 6/80	
	14. SPONSORING AGENCY CODE EPA/600/01	
15. SUPPLEMENTARY NOTES		
16. ABSTRACT This report describes an evaluation of the feasibility of using a remotely sensed data base as input into the Hydrologic Simulation Program-FORTRAN (HSPF). Remotely sensed data from the satellite LANDSAT and conventionally obtained data were used to set up the input parameters of HSPF. Simulations were run to compare the two sets of data. The remotely sensed data set performed as least as well as the conventional one when compared with observed data. In addition, it was estimated to offer savings of 30 to 50 percent in the cost of set up and operations. Areas of potential development of remote sensing for modeling input are suggested. The Image Based Information System (IBIS) used to manipulate the LANDSAT data was found to contribute significantly to the data collection and handling process.		
17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS (This Report) UNCLASSIFIED	21. NO. OF PAGES 107
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE

EPA 600/3-81-037

June 1981

EVALUATION OF REMOTE SENSING DATA
FOR INPUT INTO
HYDROLOGICAL SIMULATION
PROGRAM-FORTRAN (HSPF)

by

D.D. Franz and S.M. Lieu

Hydrocomp, Inc.
Mountain View CA 94040

Contract No. 68-01-5801

Project Officer

Thomas O. Barnwell
Technology Development and Applications Branch
Environmental Research Laboratory
Athens GA 30613

ENVIRONMENTAL RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
Athens GA 30613

PRODUCT OF:
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

Disclaimer

This report has been reviewed by the Environmental Research Laboratory, U.S. Environmental Protection Agency, Athens, Georgia and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Foreword

As environmental controls become more costly to implement and the penalties of judgement errors become more severe, environmental quality management requires more efficient analytical tools based on greater knowledge of the environmental phenomena to be managed. As part of this Laboratory's research on the occurrence, movement, transformation, impact, and control of environmental contaminants, the Technology Development and Applications Branch develops management or engineering tools to help pollution control officials achieve water quality goals through watershed management.

Computer models, such as the Hydrological Simulation Program-FORTRAN (HSPF), that simulate environmental conditions over extended periods of time have become an integral part of environmental planning and management over the last decade. To successfully model the complex and subtle components of the environment, large amounts of accurate data over wide areas are required. This report evaluates the use of the LANDSAT satellite to provide environmental data for simulation runs by HSPF on a Virginia watershed. Management of the remotely sensed data base was the responsibility of the National Aeronautics and Space Administration's Kennedy Space Center, which contributed inhouse work, contract services through the Jet Propulsion Laboratory, and partial funding through an interagency agreement with EPA. Land based data were provided by the Northern Virginia Planning District Commission. The successful completion of the project is due in large part to the contributions of these organizations.

David W. Duttweiler
Director
Environmental Research Laboratory
Athens, Georgia

Abstract

This report describes an evaluation of the feasibility of using a remotely sensed database as input into the Hydrologic Simulation Program - Fortran (HSPF). Remotely sensed data from the satellite LANDSAT and conventionally obtained data were used to set up the input parameters of HSPF. Simulations were run to compare the two sets of data.

The remotely sensed data set performed at least as well as the conventional one when compared with observed data. In addition, it was estimated to offer savings of 30% to 50% in the cost of set up and operations. Areas of potential development of remote sensing for modelling input are suggested.

The Image Based Information System (IBIS) used to manipulate the LANDSAT data was found to contribute significantly to the data collection and handling process.

This report was submitted in fulfillment of Contract No. 68-01-5801 by Hydrocomp under the sponsorship of the U.S. Environmental Protection Agency. This report covers a period from March 1979 to June 1980, and work was completed as of June 24, 1980.

Table of Contents

Disclaimer	ii
Foreword	iii
Abstract	iv
Table of Contents	v
List of Figures	vi
List of Tables	vii
Acknowledgements	ix
1 Introduction	1
2 Conclusions	2
3 Recommendations	4
4 Hydrologic and Water Quality Modelling with HSPF	5
4.1 Purpose of Modelling	5
4.2 HSPF Structure	6
5 Modelling in the Occoquan Watershed	13
5.1 The Watershed	13
5.2 Estimation of Parameters from LANDSAT imagery ...	15
5.3 Simulation Process	33
6 Summary and Analysis of Simulation Results	36
6.1 Hydrologic Results	36
6.2 Water Quality Results	36
7 Cost Effectiveness of Current and Other Studies	64
8 Potential for the Use of Remotely Sensed Data in Modelling	71
8.1 Physical Processes	72
8.2 Modelling Techniques	75
8.3 The Role of Remotely Sensed Data and IBIS Data Management	77
8.4 Methods to Assess the Utility of Remotely Sensed Information and the Use of IBIS Data Management	81
9 References	83
Appendix 1 The Image Based Information System: An Overview	84
Appendix 2 LANDSAT and IBIS Processing Steps for the Occoquan Data Base	92

List of Figures

Figure Number	Title	Page
4.1	Organization of HSPF Software	7
4.2	Conceptual View of HSPF Organization	11
5.1	The Occoquan Basin	14
5.2	Diagram of the Data Collection and Manipulation Process	16
6.1	Flow Rate at Reach 65, WY 75	52
6.2	Flow Rate at Reach 65, WY 76	53
6.3	Water Temperature at Reach 65, WY 75	54
6.4	Water Temperature at Reach 65, WY 76	55
6.5	DO Concentration at Reach 65, WY 75	56
6.6	DO Concentration at Reach 65, WY 76	57
6.7	BOD Concentration at Reach 65, WY 75	58
6.8	BOD Concentration at Reach 65, WY 76	59
6.9	NH3 Concentration at Reach 65, WY 75	60
6.10	NH3 Concentration at Reach 65, WY 76	61
6.11	PO4 Concentration at Reach 65, WY 75	62
6.12	PO4 Concentration at Reach 65, WY 76	63
8.1	Hydrologic and Water Quality Processes	73
8.2	Model Inputs, Processes and Outputs	76
8.3	Model Parameters and Key Physical Processes for HSPF	78

List of Tables

Table Number	Title	Page
5.1	Category Designations in the LANDSAT Land Use Classification	17
5.2	Aggregation of Land Use Classes for HSPF Modelling	18
5.3	Impervious Area Estimates Used by the NVPDC	19
5.4	LANDSAT Land Use Distribution Based on NVPDC Land Use Classification	20
5.5	IBIS Acreage Tabulation for Slope, by LANDSAT Land Use Classes	22
5.6	IBIS Acreage Tabulation for Total Volume of Storage Above the Restrictive Layer, by LANDSAT Land Use Classes	23
5.7	List of Model Parameters Considered for Evaluation with Remotely Sensed Database	24
5.8	Sample Calculation of Hydrologic Parameters for the Remotely Sensed Database in Subbasin 4	27
5.9	Comparison of Hydrologic Parameters of the Conventional and Remotely Sensed Databases	28
5.10	Changes in Hydrological Parameters in Subbasin 4	30
5.11	NVPDC Values for Dry Weather Pollutant Accumulation Rates for Pervious and Impervious Fractions of Urban and Rural Land Uses	31
5.12	NVPDC Values for Monthly Vegetative Cover and Sediment Potency Factors for Cropland	32
5.13	Comparison of Water Quality Parameters of the Conventional and Remotely Sensed Databases	34
5.14	Land Use Distribution by Aggregated Land Use Classes	35
6.1	Annual Per Unit Area Response of Pervious Land Surface Hydrology in Subbasin 4	37
6.2	Comparison of Monthly Outflow Volume for Subbasin 4, in inches	38
6.3	Comparison of Constituent Outflow at Reach 65	39
6.4	Comparison of Constituent Load from Each Subbasin	40
6.5	Analysis of Annual Load of Sediment and Selected Constituents by Subbasins, Water Year 1975	42

6.6	Mean Percentage of Total Water Quality Constituent Associated with Sediment Load from Pervious Land Areas, Water Year 1975	43
6.7	Details of Annual BOD Load from Land Use Segments in Subbasin 7, Water Year 1975	44
6.8	Comparison of Annual Sediment and Constituent Load from Subbasin 5, Water Year 1975	45
6.9	Detailed Comparison of Annual Load of Sediment and Selected Constituents from Subbasin 5, Water Year 1975	46
6.10	List of Symbols Used in Figs. 6.1 to 6.12	51
7.1	Original NVPDC Cost Estimates for a Conventional Database in the Ocoquan	65
7.2	Original NASA Cost Estimates for a Remotely Sensed Database in the Ocoquan	66
7.3	Estimate of Labour Rates and Other Costs	67
7.4	Comparison of Estimated Costs, Conventional versus Remotely Sensed Databases	68
8.1	Key Physical Processes in the Watershed	74

Acknowledgements

This study was conducted with the assistance and support of, in alphabetical order, the Environmental Research Laboratory of the EPA, the National Aeronautics and Space Administration (NASA), and the Northern Virginia Planning District Commission (NVPDC), under EPA contract number 68-01-5801. We would like to express our appreciation to these agencies for their cooperation, and in particular to Roy Bland of the Kennedy Space Center, Richard Fretz of the Jet Propulsion Laboratory, and Jack Hartigan of the NVPDC. We would also like to acknowledge David Kliever, who performed most of the HSPP simulations.

Section 1 Introduction

In recent years, there has been a growing awareness of the environment and the urgency of its protection. Planners and scientists alike are increasingly turning to computer modelling to assist them in the understanding of the complexity of natural phenomena and their interactions, and of the potential impact of human action upon the environment.

Yet these models are no more than numerical representations of these phenomena and interactions. To successfully model the complex and frequently subtle components of the environment, large amounts of accurate data over wide areas are required. These data are essential to the construction, execution and improvement of any model. However, what is available is often less adequate than what is desired.

The development of satellites able to remotely detect a wide variety of environmental data over large areas of the earth easily and rapidly offers tremendous potential in improving the ease and accuracy of modelling efforts. (For a partial list of some of the measurements possible, see Pabst and Ceruak, 1978.) The question is whether it is feasible to use this information from the satellites in earth orbit as input into environmental models, especially the more complex ones.

The aim of the current study is the evaluation of the issues in this use of remotely sensed data. Data from the LANDSAT satellite was supplied by the National Aeronautics and Space Administration (NASA). Terrestrially based data was provided by the Northern Virginia Planning District Commission (NVFDC). These were used as input in simulation runs with the Hydrological Simulation Program - Fortran (HSPF). The questions addressed include the feasibility of using remotely sensed data as model input, difficulties in the implementation of this approach, and the modelling accuracy and cost effectiveness of so doing in comparison with a conventionally obtained database. The contribution of ancillary computer hardware/software for handling LANDSAT information to modelling, particularly IBIS, the Image Based Information System (Bryant and Zobrist, 1977), is also assessed.

Following the Conclusions and Recommendations, Section 4 briefly describes the use of modelling, and HSPF and its input requirements. This is followed by sections on the actual use of remotely and conventionally sensed information as input for simulation runs (Section 5), and on the analyses of the results obtained (Section 6). A cost comparison of the use of the two databases and potential developments in the use of remotely sensed data in modelling are discussed in Sections 7 and 8 respectively.

Section 2 Conclusions

Conclusions may be drawn for both the results of the simulations themselves, and for the process of using remotely sensed databases in HSPF modelling.

In terms of the simulations themselves, we found that the remotely sensed database performed at least as well as the conventional database. This is supported by the following :

1. Differences in land use distribution in the Broad Run of the Occoquan Basin had little effect on the overall outcome of HSPF modelling. The changes in parameters resulting from differing land use distributions tended to balance out during simulation.
2. Both the conventional and the remotely sensed parameter sets performed well for certain simulations. The results for flow rate, water temperature and dissolved oxygen are comparable to observed data.
3. Neither parameter set seemed to perform well in nutrient simulations. However, it was difficult to assess the accuracy of these results due to the inadequacies of the observed data.
4. The parameter set from the remotely sensed database performed at least as well as the conventional set. Remote sensing also offers some additional advantages.
5. Refinement of information on urban and water areas in the remotely sensed land use classification would result in better segment definition.

In terms of using a remotely sensed database with IBIS, we conclude the following:

1. The current use of remotely sensed information is relatively indirect; land use classes are processed by a transformation function to generate modelling parameters. With the availability of the appropriate transformation function(s), the use of remotely sensed databases for HSPF and other types of modelling is highly feasible.
2. In efforts such as the current one, the use of a remotely sensed database is estimated to offer significant savings in cost (about 30% to 50%) in comparison with use of a conventional one. Much greater economies of scale are expected for larger study areas.
3. In addition to cost, a remotely sensed database also offers repeatability and relative rapidity and ease of information collection. It would be particularly advantageous in large and / or inaccessible areas, or areas requiring frequent database updates.

4. IBIS made significant contributions in verifying segment definition, and in computing basin wide average parameters and other cross tabulations. It greatly enhanced the use of the remotely sensed database. The image formatted data plane concept of IBIS is vital to the integration of various sources and types of data, particularly on a large scale.

Section 3 Recommendations

In relation to the further use of remote sensing in hydrologic and water quality simulations, we recommend the following :

1. Further development of the data collection and transformation processes as set out in the section on potential applications. The development of the latter is especially desirable since its lack is the main impediment to more intensive and extensive use of remote sensing.
2. Refinement of the remotely sensed land use classification of urban areas. This would enhance the use of satellite information in areas of rapid urban development, where updates are required to accurately assess changes in hydrology and water quality.
3. The use of remotely sensed databases for modelling agricultural or undeveloped areas, particularly if these are large or inaccessible.
4. The assessment of information which can be obtained for modelling purposes from future satellite developments. In longer terms, this is a valuable avenue of research, since the availability of new data may significantly affect or even entirely modify the modelling approach.
5. The utilization of IBIS occurred subsequent to the initiation of the current effort. We believe that a more comprehensive evaluation of the use of HSPF and Remote Sensing, particularly with the use of IBIS, should be conducted on a significantly larger area. In conjunction, it is recommended that the digital terrain data from the Defence Mapping Agency (DMA) be utilized for improving basin slope determinations.

Section 4 Hydrologic and Water Quality Modelling with HSPF

Digital computer models, simplified mathematical representations of various phenomena, have become an integral part of planning and management processes in the last ten years. Many different kinds of models have been devised to meet the needs of individual situations.

In resource management, particularly water resources management, models must be able to deal with complex interactions that may extend over long periods of time, where the frequency and probability of occurrence of an event may be of vital importance. Simulation of a "critical period" cannot provide the information necessary to determine the probability of hydrologic and water quality events. Planners need a method to determine the frequency characteristics of these events within a watershed under alternative plans for its future development. This need led to the development of models which can simulate extended time periods, from five to more than a hundred years, with varying meteorologic and flow conditions (continuous dynamic simulations). Often, the performance of a proposed project or management scheme is then inferred from these simulated time series.

Section 4.1 Purpose of Modelling

The development of models with these capabilities, such as HSPF, serves three major purposes.

1. Understanding the real world phenomena

The development of a computer model synthesizes what is known about the physical system and defines areas where there is a lack of rigorous information. In so doing, various components and their interactions must be defined. This process of definition often forces a synthesis of information which ultimately yields a clearer understanding of the processes. Quantitative relationships must be developed or assumed to construct the model. Application of the model then provides estimates of the accuracy of these relationships. Thus modelling contributes to the understanding of the physical world itself.

2. Extension of time series

Long records of historic hydrologic and water quality data are preferred to reliably establish the frequency of critical events. In most situations, such data are not available. Moreover, the effects of urban development and other watershed changes can bias a long historic record. Continuous simulation can be used to develop long records of hydrology and water quality for existing as well as historic or planned conditions. For example, a simulation model can be adjusted (calibrated) to represent the watershed characteristics by using a short streamflow record (of length 3 to 5 years) with little change in conditions. Long term meteorological

logic data are then input to the model to obtain a simulated record of sufficient length for the generation of a streamflow frequency curve.

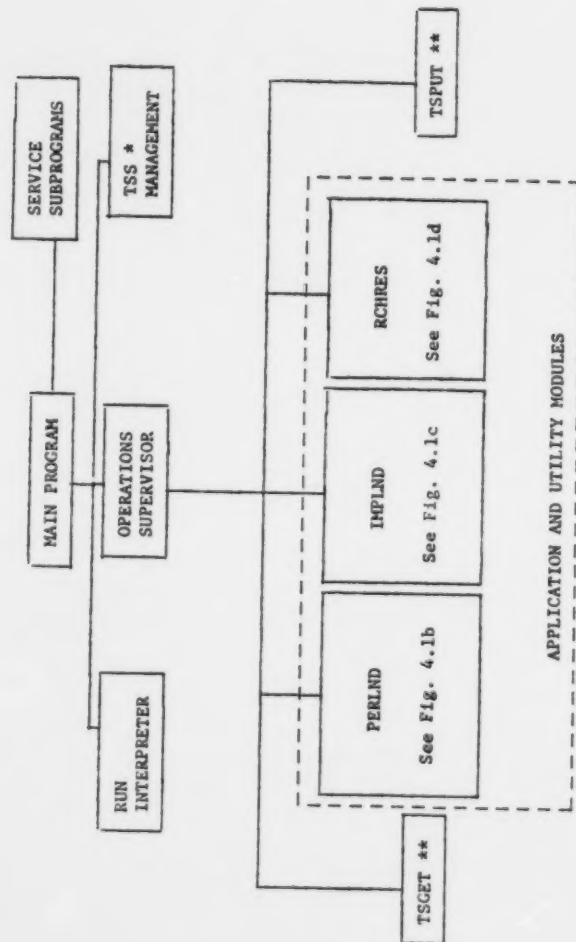
3. Prediction

In water resource planning, it is often necessary to consider a number of management alternatives where the results cannot be assessed intuitively. A continuous dynamic model, calibrated for the area in question, can be used to predict the impact of these different alternatives. This aids in the selection of the most desirable plan as well as the establishment of new and better ones. This is the most obvious purpose of modelling although, as shown above, it is not the only one.

Section 4.2 HSPF Structure

HSPF was designed to simulate hydrologic and water quality processes in natural and man-made water systems. It is organized into modules, each simulating a number of related hydrologic or water quality processes (Fig. 4.1). These modules, small models within the larger whole, may be activated or omitted by choice. Each module is further subdivided into sections and subsections which contain the actual algorithms required to represent particular interactions in the water system. An example is the algorithms for pesticide adsorption onto and desorption from soil particles. These are found in the section PEST, which deals with all pesticide interactions that occur on pervious land. The section PEST is within the module PERLND, which simulates all the hydrologic and water quality processes on pervious land. Details on HSPF may be found in the HSPF executive summary (Grimsrud, Franz, Johanson and Crawford, 1979) and user's manual (Johanson, Imhoff and Davis, 1979).

A water system may be a watershed or it can equally be something artificial, for example a canal network. For simplicity, it is generally referred to as a watershed. The characteristics of the watershed and its environment are input into the model either as constant parameters, or as time series, i.e., quantities that are variable over time. (See Fig. 4.2.) Examples of time series include air temperature, precipitation, solar radiation and evapotranspiration. Constant inputs include parameters representing watershed characteristics such as overland flow slope, channel slope, reaeration coefficients, nutrient reaction parameters, etc. Another group of inputs is the state variables. These represent the "initial conditions" of certain features of the watershed at the start of the simulation, and may be modified over the course of the simulation. Examples of these include soil moisture storage, snow depth and ground water storage. Sometimes, the information available is not directly useable in HSPF. In these cases, a translator or "preprocessor" is required to convert the data available into an acceptable form.



* TSS = Time Series Store
 ** TS = Time Series

Figure 4.1a Organization of HSPF Software

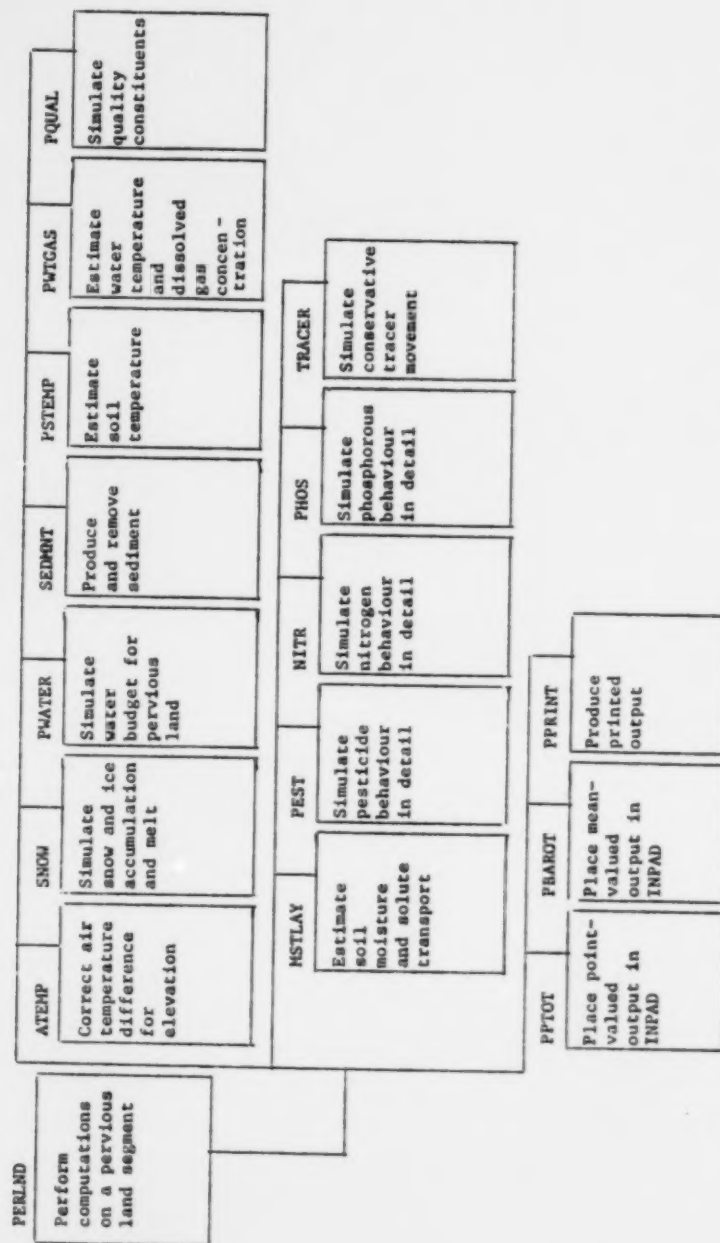


Figure 4.1b Subsections Within the Module PERLND

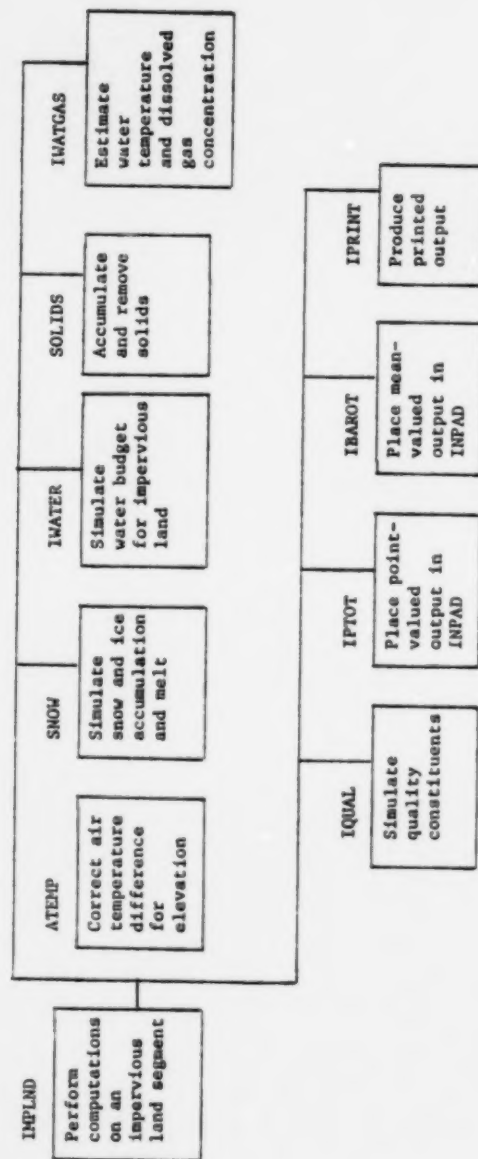


Figure 4.1c Subsections Within the Module IMPLND

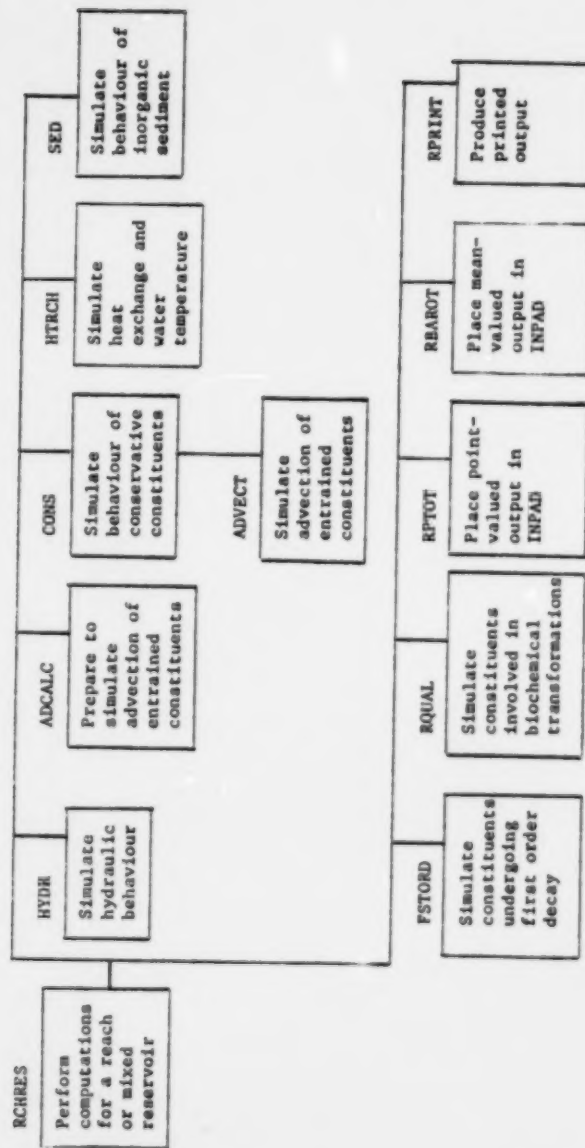


Figure 4.1d Subsections within the Module RCHRES

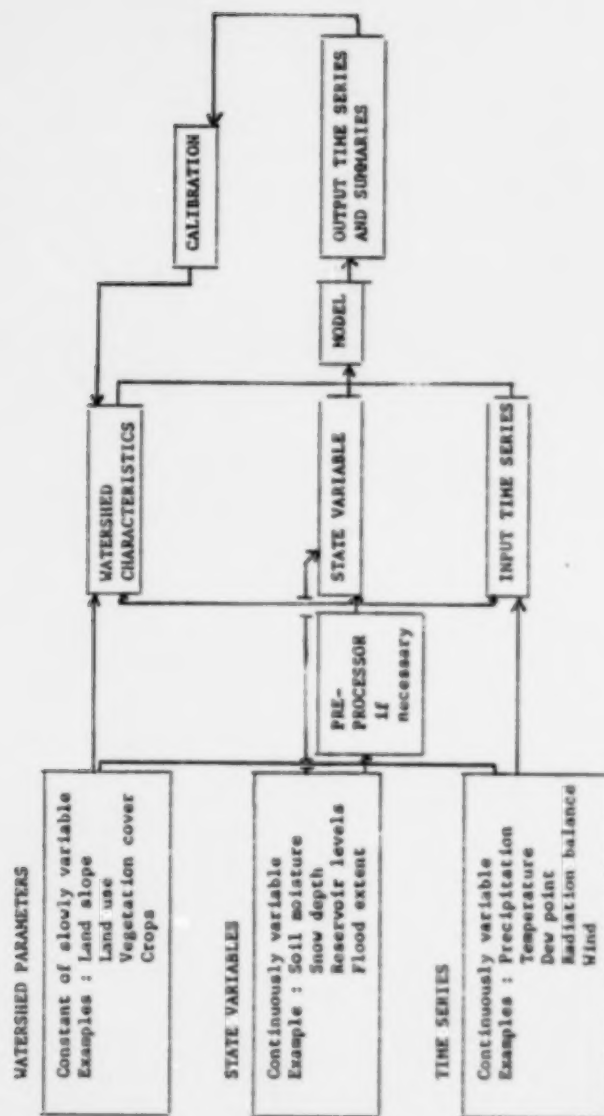


Figure 4.2 Conceptual View of HSPF Organization

All the interactions are calculated by the relevant algorithms for each user defined time interval (ranging from one minute to one day), and are repeated over the number of steps within the defined period of simulation.

As output, HSPF produces a time series of values for each of the simulated conditions, e.g. flow rate, water temperature or water quality constituent loading. These time series may then be compared with observed values. The model is calibrated to the observed data by adjusting the constant watershed parameters in the model input to reduce the discrepancies between the simulated and the observed behaviour of the watershed.

Section 5.0 Modelling in the Occoquan Watershed

The Occoquan Basin was selected for this study because extensive simulation of this watershed had been done by the NVPDC. As a result, watershed parameters, initial conditions and time series derived by conventional means were available. Thus the results using this conventional data could be compared with results obtained using a parameter set estimated in part from land use distribution inferred from LANDSAT images.

In the current study, the data obtained thorough conventional means by the NVPDC, including the land use distribution of the basin, was supplied by the agency. A remotely sensed land use distribution obtained through LANDSAT imagery was supplied by the Kennedy Space Center and a database relating the terrestrially derived information to the Landsat land classifications was provided by the Jet Propulsion Lab (JPL). It should be mentioned that in the following discussions and comparisons of parameters and simulation results, the conventional and remotely sensed data sets are often referred to, for brevity, as "NVPDC" and "LANDSAT" respectively. Beyond the provision of parameters and land use distributions, neither was involved in the actual HSPF simulation runs and the results so obtained.

Section 5.1 The Watershed

The Occoquan is a watershed about 580 square miles in size near Fairfax, Virginia. It has been divided into 15 subbasins by the NVPDC for previously conducted studies. In the current study, only subbasins 4, 5 and a part of subbasin 7 were modelled. These areas include the Broad Run Reservoir, and the Broad Run River and its tributaries up to the sampling station above Lake Jackson (Fig. 5.1), for a total of about 95 square miles. In the simulations with HSPF, the land areas are called subbasins 4, 5 and 7 in keeping with the original system, although only part of subbasin 7 was actually involved. The channels associated with each area are defined as Reach 40, 50 and 65 respectively. Thus the sampling point is at the outflow of Reach 65.

The NVPDC had previously used specially modified versions of the Non Point Source (NPS) and HSPX QUALITY models for simulations of the Occoquan (Hydrocomp, Inc., 1978). HSPF incorporates the capabilities of these models but contain minor changes in parameter definition. As part of the task of testing HSPF for the EPA, the set of parameters developed by NVPDC for the Occoquan had been adapted for HSPF by Hydrocomp. Comparison of HSPF and NVPDC simulations verified that HSPF produced results which were very close to those of NPS and HSPX QUALITY. Details of this comparison and the sources of time series data are given in the final report on the maintenance and testing of HSPF (Johanson and Kliever, 1980).



Figure 5.1
The Occoquan Basin

Section 5.2 Estimation of Parameters from LANDSAT Imagery

Selected parameters were estimated using information supplied by the Kennedy Space Center and JPL. The following sections document the nature of the information available and the methods used to estimate the parameters. A diagram showing all the data collection and manipulation processes can be found in Fig. 5.2.

Section 5.2.1 Nature of Information

The land use information supplied to Hydrocomp was derived from LANDSAT imagery using the IMAGE 100 system. Image 100 is a hardware and software system that processes spectral information obtained by LANDSAT to provide the more direct information required. A composite of two days, LANDSAT bands 6 and 7 from 8th April, 1975 and bands 4 and 5 from 3rd August, 1975 were combined to define a composite information base for the classification of land use. Much effort was expended on the development of a wide range of land use classes, but difficulties in discriminating between some of these were encountered. The final classification consists of eight categories as well as an "unclassified" one. The eight category designations used in this classification are shown in Table 5.1. Approximately 400 ground truth points were established by field observation. A classification accuracy of LANDSAT analysis to ground truth was about 70%.

The land use classification was the only information available from LANDSAT imagery. However, it was possible to relate this remotely sensed land use to an extensive digital database of terrestrially sensed information developed by the NVPDC through the use of the Image Based Information System (IBIS) of the Jet Propulsion Lab. Appendix 1 contains a brief summary of this system.

The NVPDC had compiled the terrestrially sensed database using a spatial information system called GRIDS developed at Harvard University, (Roffe, 1977). The GRIDS information was averaged over a square 1,000 feet on a side whereas the LANDSAT information was averaged on a square 79 meters on a side. IBIS was able to disaggregate the GRIDS information to match the resolution of the LANDSAT.

Also, the LANDSAT land use distribution was converted into the NVPDC classification to facilitate parameter calculation. Table 5.2 shows the corresponding classes between the two databases. In so doing, unclassified areas in the LANDSAT system were distributed evenly throughout the other classes. The "water" class was also distributed in subbasins 4 and 7; however, in subbasin 5, this was aggregated with the "industrial" class to represent pollutant input into the Broad Run Reservoir as a result of precipitation. As a result of this reclassification, the water areas were assigned other characteristics from the GRIDS database. In addition, the pervious to impervious acreage ratio estimated by the NVPDC for each land use class was adhered to in the remotely sensed database (Table 5.3). Table 5.4 gives the LANDSAT land use distribution for the NVPDC land use classes.

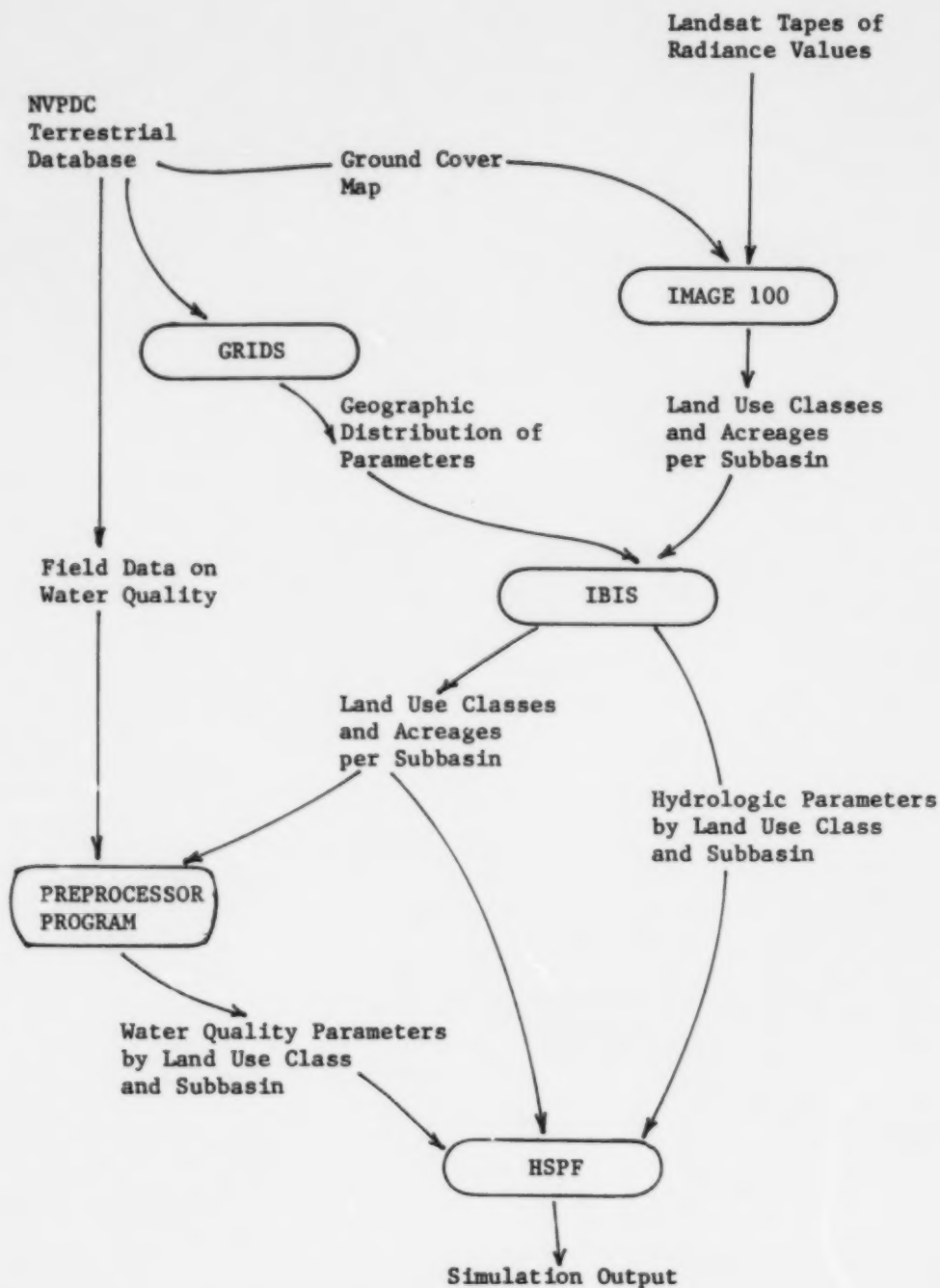


Figure 5.2 Simplified Diagram of the Data Collection and Manipulation Process

Table 5.1 CATEGORY DESIGNATIONS IN THE LANDSAT
LAND USE CLASSIFICATION

Forest	: Forest
Und/Clr	: Undeveloped or Cleared
Com/Ind	: Commercial or Industrial
Hvy Res	: Heavy Residential
Med Res	: Medium Residential
Water	: Water Surfaces
Pasture	: Pasture
Grains	: Agricultural Land in Various Grains

Table 5.2 AGGREGATION OF LAND USE CLASSES FOR HSPF MODELING

HSPF CLASSIFICATION	NVPDC CLASSIFICATION	LANDSAT CLASSIFICATION*
0 Forest/Idle	Forest ↔ Forest Idle ↔ Undeveloped/Cleared	
1 Agricultural Cropland	Cropland ↔ Grains	
2 Agricultural Pasture	Pasture ↔ Pasture	
3 Urban Residential	Estate single family Large lot single family Medium density single family ↔ Medium residential Townhouse ↔ Heavy residential Highrise	
4 Urban Employment	Industrial ↔ Commercial & Industrial Institutional Shopping Center Central business	

↔ indicates correspondence of classes in the computation of parameters.

- * Unclassified areas distributed evenly throughout.
Water areas distributed evenly throughout subbasins 4 and 7, aggregated with Urban Employment in subbasin 5 (Broad Run Reservoir).

TABLE 5.3 IMPERVIOUS AREA ESTIMATES USED BY THE NVPDC

<u>Land Use</u>	<u>Impervious Area/Total Area (%)</u>
Non-Urban*	1
Large-Lot Single Family	12
Medium Density Single Family**	25
Townhouse/Garden Apartments	40
Commercial	90
Industrial***	70
Institutional (e.g., churches, schools)	35

* Corresponds to the forest/idle, agricultural cropland and agricultural pasture classes in the aggregated classification.

** Corresponds to the urban residential class in the aggregated classification. Includes only impervious cover directly connected to a paved or improved drainageway.

*** Corresponds to the urban employment class in the aggregated classification.

TABLE 5.4 LANDSAT LAND USE DISTRIBUTION BASED ON NVPDC LAND USE CLASSIFICATION

NVPDC Land Use* Classes	Forest (Forest)	Cropland (Grains)	Pasture (Pasture)	Idle (Und/Clr)	Industrial (Comm/Ind)	Townhouse/ Garden Apts. (Hvy Res)	Mod-Density Single Family (Med Res)
Subbasin							
4	15500	3707	6841	5031	550	61	886
5	7567	1517	2174	1658	312	50	640
7	12551	2853	3828	2937	992	161	1099

* Corresponding LANDSAT classes in parenthesis.

Details of the LANDSAT and IBIS processing steps are included in Appendix 2.

The information on slope and total volume of storage above the restrictive layer in Tables 5.5 and 5.6 respectively are examples of information taken from the GRIDS database and processed by IBIS. With its resolution, the GRIDS information did not isolate water bodies (though LANDSAT imagery also misclassified some of these). Thus, The non-zero slope for water surfaces and the very small storage for the water class in these tables arise from differences in resolution between GRIDS and LANDSAT.

Also, the initial IBIS images produced from the slope information taken from GRIDS revealed an apparent discontinuity in slope along a north-south line traversing subbasin 4. Subsequent examination of the base topographic maps from which the GRIDS data on slope was estimated revealed that the contour interval increased by a factor of two across this line. This area corresponded generally to a highly forested mountain area, Pignut Mountain. The increase was apparently not noticed by the original analyst, so that slopes in the GRIDS system for points west of this line were one half of their true value. IBIS was able to apply an approximate correction using a transformation function developed from the original GRIDS discretization scheme. Both the detection and correction of this discrepancy in land surface slope demonstrated the usefulness of IBIS in processing spatially indexed information.

Section 5.2.2 Parameters Estimated

Table 5.7 shows the list of parameters suggested by the project contract to be quantifiable based on remotely sensed information. These are listed under the parameters used in the NPS and HSPX models, and arranged in order of the feasibility of being measured by remote sensing. It should be pointed out that the first entry in Table 5.7, Area is not really a parameter as the term is used here. Although the area of a segment under simulation is vital to the results produced, it is used as a scaling factor in the organization of the simulation, area (and net result), and not as a constant term characterizing the simulation itself. Therefore it is not referred to as a parameter in this discussion.

In the current study, LANDSAT imagery provided only information on land use classes per subbasin and the area of each class. In order to derive parameters from this information, correlations between land use and parameters must be established. This was achieved for three of the parameters from the list of 20 in Table 5.7 through previous work by the NVPDC through the use of IBIS. (These are marked with an asterisk (*) in the table.) Other parameters in this list could not be defined from the remotely sensed land use information for the following reasons:

TABLE 5.5 IBIS ACREAGE TABULATION FOR SLOPE, BY LANDSAT LAND USE CLASSES

Sub-basin	Land Use	Mean Slope (%)	Acreage
4	Unclass	10.4	2811.4
4	Forest	13.6	14161.8
4	Und/clr	8.6	4597.2
4	Com/ind	9.7	502.7
4	Hvy res	6.8	55.5
4	Med res	8.8	809.6
4	Water	7.8	103.3
4	Pasture	7.7	6250.4
4	Grains	7.7	3386.6
4	Entire subbasin	10.7	32678.7
5	Unclass	4.7	903.7
5	Forest	5.1	7075.5
5	Und/clr	4.7	1549.9
5	Com/ind	4.6	291.5
5	Hvy res	4.7	46.3
5	Med res	4.4	593.4
5	Water	4.6	555.2
5	Pasture	4.3	2032.6
5	Grains	4.7	1418.8
5	Entire subbasin	4.8	14471.8
7	Unclass	3.8	1981.7
7	Forest	4.6	11532.4
7	Und/clr	4.3	2698.8
7	Com/ind	3.5	911.4
7	Hvy res	3.4	148.0
7	Med res	3.7	1010.1
7	Water	5.2	54.0
7	Pasture	3.9	3517.7
7	Grains	4.1	2621.7
7	Entire subbasin	4.3	24475.9

TABLE 5.6 IBIS ACREAGE TABULATION FOR TOTAL VOLUME OF STORAGE
ABOVE RESTRICTIVE LAYER, BY LANDSAT LAND USE CLASSES

Sub-basin	Land Use	Mean Volume (in.)	Acres
4	Variable	7.0	7040.0
4	Unclass	10.7	2193.0
4	Forest	10.8	9424.2
4	Und/cir	10.6	4028.2
4	Com/ind	10.0	396.3
4	Hvy res	10.0	55.5
4	Med res	10.2	697.1
4	Water	8.9	80.2
4	Pasture	10.5	5706.1
4	Grains	10.2	3058.1
4	Entire subbasin	9.8	32678.7
5	Unclass	8.2	852.8
5	Forest	8.0	6790.2
5	Und/cir	8.6	1469.7
5	Com/ind	8.0	279.1
5	Hvy res	8.3	43.2
5	Med res	7.5	562.9
5	Water	8.2	532.1
5	Pasture	8.6	1884.5
5	Grains	8.1	1330.9
5	Entire subbasin	8.2	13745.4
7	Unclass	7.2	1950.9
7	Forest	7.1	11065.1
7	Und/cir	7.5	2680.3
7	Com/ind	7.0	908.3
7	Hvy res	7.8	146.5
7	Med res	7.2	997.8
7	Water	6.0	46.3
7	Pasture	7.6	3508.5
7	Grains	7.5	2604.7
7	Entire subbasin	7.3	23908.3

TABLE 5.7 LIST OF MODEL PARAMETERS CONSIDERED FOR EVALUATION
WITH REMOTELY SENSED DATA

Parameter	Remote Sensing Measurements	Feasibility***	HSPF Present Data Base
Area	Area	1	Topo/photos
A	Land use	1	Topo
F	Veg class	1	Topo
TIMTIL	Tillage	1	Farm records
* NN	Veg class	2	Handbooks
PETMUL	Surface temp/albedo	2	Pan evap+calib
* K3, K24EL	Veg class/density	2	Exp/crop
KUGI	Forest class	2	Inspect/exp
COVPMO	Veg class/density**	2	Tables/exp
TIMHRV	Veg class	2	Farm records
* EXPM	Veg class/density**	3	Exp/crop growth
KSER	NN-veg class	3	Exp/tables
NAPPL	Residue**	3	Farm records
ULUPTK, LZUPTK	Crop class **	3	Records/exp
MPAK, PACK	Snow area/depth	4	Exp/calib
KPL	Crop class (temp)	4	Records/exp
PLNT-N, PLNT-P	KPL-plant temp	4	Records/exp

* Evaluated in this study

** Monthly database

*** Feasibility: 1 = current capability can accomplish
2 = possibility of current capability
3 = possible with near term technology development
4 = possible with future technology

1. Not relevant to the simulations being run.

F, KUGI, MPAK and PACK.

These are parameters relating to snow accumulation and melt. Snow was not simulated by the NVPDC study and these were excluded.

TIMHRV, NAPPL, ULUPTK, LZUPTK, KPL, PLNT-N and PLNT-K.

These are parameters from the Agricultural Runoff Management (ARM) model. Since NVPDC used only the NPS and HSPX QUALITY models, these parameters could not be evaluated for lack of a baseline simulation.

K24EL.

This is the percentage of land with groundwater evapotranspiration at the full rate. It is mostly related to marshland areas which are of insignificant acreage in the area under study.

2. Relevant remote sensing information was not available.

TIMTIL.

This parameter relates to the time of tillage of agricultural land. The remotely sensed data for this project was available at only one point in time, so that the time of tillage could not be estimated.

L and A.

The overland flow length, L, could not be estimated from the remotely sensed data since this requires topographic data or at least detailed stream network information. Neither of these, nor the percentage of impervious land area (A), was supplied by the land use information from LANDSAT.

COVPMO.

This parameter relates to the variation of crop cover over the growing season. Again, although this can be obtained through remote sensing, information over time was not available to this study.

PETHUL.

This relates to temperature and albedo, neither of which was available from the remotely sensed data.

3. Relationship between land use and parameter value not known.

KSER.

This relates primarily to detailed surface roughness and is normally determined through calibration.

Other hydrologic parameters not on this list were also cross tabulated by land use classes by means of IBIS. Examples of these include subbasin slope, interflow and the volume of storage above the restrictive layer.

It was further possible to correlate land use classes with water quality parameters through relationships established in field studies by the NVPDC. These include loading of sediment and sediment associated pollutants from the land surface. Both hydrologic and water quality parameters used in this study are discussed in detail in the following sections (5.2.3 and 5.2.4).

Section 5.2.3 Hydrology Parameters

The application of HSPF to a watershed to simulate the hydrologic response requires that the watershed be subdivided into segments which can be considered homogeneous in their meteorologic and hydrologic characteristics. Naturally this requirement is only approximately met in any application. The original studies in the Ocoquan resulted in the 15 subbasins shown in Fig. 5.1 and thirteen land use classes. The latter were aggregated into five classes for computer simulation in the original NVPDC studies (Table 5.2).

Since IBIS is capable of developing many cross-tabulations of spatially indexed characteristics, one major hypothesis was that perhaps segments should be defined by land use rather than by subbasins, resulting in five segments for both hydrology and water quality. Such a redefinition of the watershed would be based on land use classes being more nearly homogeneous with respect to characteristics such as soil texture, land slope, etc. than the subbasins. Examination of the information in Tables 5.5 and 5.6 did not support this hypothesis. Land use, as defined by LANDSAT, did not have a significant correlation with any of these hydrologic characteristics. Variations of hydrologic parameters were greater between basins for the same land use class than were variations between land use classes within a basin. Therefore, the original watershed representation was retained for the simulation based on remotely sensed data.

New subbasin average values for parameters NN(overland flow roughness), K3(index to the root depth in the segment) and EPXM(maximum interception storage) were computed for each of the subbasins. The process of the computation is illustrated in Table 5.8. In each case a transformation function between land use classes and values of the parameters developed by the NVPDC was used. The values of the parameters for the entire subbasins were then calculated on an area weighted basis. Table 5.9 summarizes the results for the other subbasins. Clearly these parameters are not significantly affected by the differences in land use distribution.

TABLE 5.8 SAMPLE CALCULATION OF HYDROLOGIC PARAMETERS FOR THE REMOTELY SENSED DATABASE IN SUBBASIN

Land Use	Mean Parameter Value			Land Use Area (acres)
	Overland Flow Roughness	Lower Zone Evapo-transpiration	Maximum Interception Storage	
	(NN/NSUR)*	(K3/LZETP)*	(EXPM/CEPSC)*	
Forest	.35	.50	.15	15548
Undeveloped/Cleared	.30	.40	.12	5047
Pasture and grains	.25	.30	.10	10581
Residential and commercial/industrial	.25	.40	.05	1502
Entire subbasin	.31	.42	.12	32678
Baseline basin-wide average	.30	.43	.12	---

* HSPX and HSPF parameter names respectively.

** Subbasin wide mean defined as $\frac{\sum_{i=1}^4 A_i(X_i)}{\sum_{i=1}^4 A_i}$ where A is the area in acres

and X_i is the parameter value

for the i th land use class. For example, for the overland flow roughness NN, the mean = $15548(.35) + 5047(.30) + 10581(.25) + 1502(.25)$

32678

TABLE 5.9 COMPARISON OF HYDROLOGIC PARAMETERS OF THE CONVENTIONAL
(NVPDC) AND REMOTELY SENSED DATABASES

Parameter	Subbasin 5		Subbasin 7	
	NVPDC	LANDSAT	NVPDC	LANDSAT
Overland flow roughness (NN/NSUR)*	.30	.31	.30	.31
Lower zone evapotranspiration (K3/LZETP)*	.42	.43	.43	.42
Maximum interception storage (EIPM/CEPSC)*	.12	.12	.12	.12

*HSPX and HSPF parameter names respectively.

The error in land slope in subbasin 4 mentioned in Section 5.2.1 also affected some parameters. The mean slope relates directly to the overland flow slope parameter in HSPF. A further, approximate, relationship was developed between average basin slope and the interflow parameter for the previous studies. This relationship was based on calibration results from three subbasins, including subbasin 4. Strictly speaking the relationship was no longer valid since the slope used in subbasin 4 was in error. However, the relationship was used to predict a new value of the interflow parameter as if the relationship were valid and subbasin 4 had not been used to develop the relationship. Table 5.10 shows the resultant changes in these two parameters for subbasin 4.

Section 5.2.4 Water Quality Parameters

The water quality parameters were again estimated using transformation functions relating land use class to pollutant loading rates and potency factors. The potency factor is defined as the number of pounds of pollutant per ton of sediment in simulations with HSPF. (The potency factors for the Non Point Source model used by NVPDC was expressed as pounds of pollutant per 100 pounds of sediment). The loading rates and potency factors for the land uses in the Occoquan were established by the NVPDC after an extensive series of small area measurements was made (Regional Resources Division, NVPDC, 1979). These are shown in Tables 5.11 and 5.12. It should be noted that the potency factor is implicit in Table 5.11 as the appropriately scaled ratio between sediment and the other constituents.

The water quality simulations done by NVPDC did not utilize the detail contained in the thirteen different land use classes. Instead, these were aggregated to five classes: 0, Forest/idle; 1, Agricultural cropping; 2, Agricultural pasture; 3, Urban residential; and 4, Urban employment. Some of these classes did not involve any aggregation e.g., agricultural-pasture, whereas others did, e.g. forest/idle. The area weighted average of the loading rates and potency factors were computed for those land uses which required aggregation. This detail is important in understanding the possible parameter changes to be expected from the land use changes indicated by the LANDSAT imagery. Most of the parameters affect the simulated response of a unit area. (In fact the only parameter which does not have a unit area effect is the total area of a given land use.) The calculation of per unit area parameters for the aggregated land use classes is an area weighted process dependent on the acreages involved in the component classes. For example, the loading parameters of the forest/idle class change whenever the area of either forest or idle land changes. The parameter changes per unit are then compounded because the load on the stream is computed for the entire land use area, i.e., it is the product of the per unit area load and the area of the aggregate land use itself.

The NVPDC had developed a computer program which incorporated the information contained in Tables 5.11 and 5.12. This obviated the need for the tedious computations required to compute the area weighted averages for the water quality parameters. Selected parameter changes resulting from the

Table 5.10 CHANGES IN HYDROLOGICAL PARAMETERS IN
SUBBASIN 4

Parameter	Original Value	Revised Value
Land Slope	7.6%	10.7%
INTFW	1.95	2.26

Table 5.11 NVPDC VALUES FOR DRY WEATHER POLLUTANT ACCUMULATION RATES (LBS/ACRE/DAY) FOR PERVIOUS AND IMPERVIOUS FRACTIONS OF URBAN AND RURAL LAND USES

	SEDIMENT		BOD		TOTAL P		TOTAL N		LEAD		ZINC	
	IMPERV	PERV	IMPERV	PERV	IMPERV	PERV	IMPERV	PERV	IMPERV	PERV	IMPERV	PERV
Estate Single Family (0.05-0.2 DU/Ac)**	2.2	1.2	0.13	0.07	0.004	0.0014	0.04	0.011	0.003	0.0002	0.003	0.0003
Large-Lot Single Family (0.5-2.0 DU/Ac)	5.5	1.0	0.20	0.13	0.01	0.0035	0.08	0.02	0.005	0.0002	0.004	0.0008
Medium Density Single Family (3-6 DU/Ac)	5.5	2.0	0.20	0.26	0.01	0.0070	0.08	0.04	0.007	0.0028	0.004	0.0016
Townhouse/Garden Apartment (6-20 DU/Ac)	5.5	2.0	0.20	0.26	0.01	0.0070	0.08	0.04	0.007	0.0028	0.004	0.0016
High-rise Residential (>30 DU/Ac)	3.5	0.7	0.63	0.07	0.006	0.0017	0.05	0.011	0.008	0.0014	0.004	0.0006
Industrial	2.5	0.7	0.63	0.07	0.006	0.0017	0.05	0.011	0.01	0.0014	0.008	0.0006
Institutional	2.5	0.7	0.35	0.07	0.006	0.0017	0.05	0.011	0.007	0.0009	0.004	0.0004
Shopping Center	2.5	0.7	0.63	0.07	0.006	0.0017	0.05	0.011	0.01	0.0014	0.008	0.0006
Central Business District	2.5	0.7	0.76	0.07	0.01	0.0017	0.09	0.011	0.02	0.0014	0.01	0.0006
Pasture	0.5	1.5	0.09	0.35	0.0013	0.0047	0.015	0.047	0.0025	0.0002	0.002	0.0008
Forest*	0.5	1.0	0.09	0.036	0.0013	0.0004	0.015	0.006	0.0025	0.0001	0.002	0.0001
		2.0		0.048		0.0005		0.0084		0.0002		0.0002
Idle	0.5	1.0	0.09	0.034	0.0013	0.0005	0.015	0.008	0.0025	0.0001	0.002	0.0002
Cropland (Conventional Tillage)	5.8											
Cropland (Minimum Tillage)	0.0											

SEE TABLE 5.12

SEE TABLE 5.12

*For pervious fraction accumulation rates, top value is recommended for January through September and bottom value is recommended for October through December.

** DU = Dwelling unit

TABLE 5.12 NVPOC VALUES FOR MONTHLY VEGETATIVE COVER AND SEDIMENT POTENCY FACTORS
FOR CROPLAND

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(a) Vegetative Cover (# of total area)												
Conventional Tillage ¹	-	-	-	-	20	50	75	90	70	15	-	-
Minimum Tillage	40	40	60	70	90	90	90	90	70	30	40	40
(b) BOD Potency Factors (lb/100 lb)												
Conventional Tillage	0.8	0.8	1.0	1.0	1.0	1.3	1.5	1.5	1.5	0.8	0.8	0.8
Minimum Tillage	0.9	0.9	0.9	1.1	1.4	1.4	1.4	1.4	1.4	0.9	0.9	0.9
(c) Total P Potency Factors (lb/100 lb)												
Conventional Tillage	0.08	0.08	0.10	0.10	0.10	0.13	0.15	0.15	0.15	0.08	0.08	0.08
Minimum Tillage	0.09	0.09	0.09	0.11	0.14	0.14	0.14	0.14	0.14	0.09	0.09	0.09
(d) Total N Potency Factors (lb/100 lb)												
Conventional Tillage	0.32	0.32	0.40	0.40	0.40	0.52	0.60	0.60	0.60	0.32	0.32	0.32
Minimum Tillage	0.36	0.36	0.36	0.44	0.56	0.56	0.60	0.60	0.60	0.36	0.36	0.36
(e) Lead Potency Factors (lb/100 lb)												
Conventional Tillage	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Minimum Tillage	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
(f) Zinc Potency Factors (lb/100 lb)												
Conventional Tillage	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
Minimum Tillage	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

¹ To simulate the ground disturbance which occurs during tillage periods, the amount of sediment available for washoff is increased by 1.0 ton/acre during the second week in May and the final week in September.

LANDSAT land use are shown in Table 5.13.

For the most part the differences in the per unit area parameters were small though the areas for a given aggregate land use were greatly changed in many cases. Table 5.14 shows the land use distribution for both LANDSAT and the NVPDC values. The differences were marked in subbasin 4 and the preliminary results caused the NVPDC to re-evaluate their estimates. The original NVPDC land use estimates in subbasin 4 were made quickly and were based primarily on available farm records because subbasin 4 was outside the authorized 208 planning area so that funds for land use surveys were limited. (Also, the NVPDC focussed its data collection effort downstream of the reservoir in subbasin 5, since this reservoir tended to act as a buffer for upstream inputs.) The re-evaluation utilized aerial photographs as a check on farm records and showed that apparently the farm records exaggerated the acreages under agriculture. The revised NVPDC values for subbasin 4 (Table 5.13) reflect approximately the same degree of agreement with the LANDSAT values as in subbasins 5 and 7.

Section 5.3 Simulation Process

The simulation process began once the parameters had been estimated. This was carried out for the water years 1975 and 1976, where meteorological data were most complete. The major steps in the simulation process are as follows:

1. Simulate the per unit area hydrologic response of each subbasin and store the results for access by later steps of the process.
2. Simulate the per unit area production of water quality constituent for each of the five land uses on the three subbasins. Compute the aggregate loading for each constituent by the product of land use area and per unit area production. Store these values for access by later steps of the process.
3. Simulate the response of the stream to the aggregate loadings computed in step 2. Store selected results for later analysis.

These three steps were taken for both the set of parameters developed by NVPDC by conventional means and for the set of parameters partially re-estimated using LANDSAT imagery for the following reason. The aim of the modelling study was the comparison of HSPF simulations based on parameters from the NVPDC against those based on remotely sensed information. Because of the small differences that were found in the comparison between HSPF and NPS/HSPX QUALITY results (Johanson and Kliever, 1980), as would be expected in any comparison of two large model simulations, a baseline HSPF simulation must be made. Without this, changes in output caused by parameter changes using remotely sensed data may be masked by the small changes due to the model differences themselves.

Table 5.13 COMPARISON OF WATER QUALITY PARAMETERS OF THE CONVENTIONAL AND REMOTELY SENSED DATABASES

SUBBASIN/LANDUSE*	PARAMETER	MONTH	UNITS	NVPDC	LANDSAT	% DIFFERENCE**
4 0	Vegetation Cover	Oct-Mar	%	.900	.920	+ 2.2
5 0				.915	.916	- 0.3
7 0				.915	.917	+ 0.2
4 0	Vertical Sediment Input	Oct-Dec	lb/ac.d	1.48	1.30	-12.0
5 0				1.32	1.35	+ 2.3
7 0				1.35	1.34	- 0.7
4 3		Jan-Dec		0.74	0.78	+ 5.4
5 3				0.74	0.78	+ 5.4
4 0	Potency Factor, Organic -N	Jan-Mar	lb/ton	9.40	10.80	+14.9
5 0				10.66	10.44	- 2.6
7 0				10.33	10.50	+ 1.9
4 0		Apr-Sep		11.20	12.20	+ 8.9
5 0				12.06	11.92	- 1.2
7 0				11.88	11.96	+ 0.7
4 0		Oct-Dec		7.60	9.50	+25.0
5 0				9.24	8.98	- 2.8
7 0				8.90	9.04	+ 1.6
4 0	Potency Factor, NO ₂ , NO ₃	Jan-Mar	lb/ton	0.48	0.50	+ 4.2
4 0		Oct-Dec		0.40	0.44	+10.0
4 0	Potency Factor, Organic P	Jan-Mar	lb/ton	0.60	0.70	+16.7
7 0				0.66	0.68	+ 3.0
4 0		Apr-Sep		0.80	0.84	+ 5.0
4 0		Oct-Dec		0.40	0.54	+35.0
5 0				0.52	0.50	- 3.9
7 0				0.50	0.52	+ 4.0
4 0	Potency Factor, BOD	Jan-Mar	lb/ton	60.0	62.0	+ 3.3
5 0				61.7	61.4	- 0.5
7 0				61.4	61.5	+ 0.3
4 0		Apr-Sep		72.0	71.0	- 1.4
5 0				71.2	71.3	+ 0.2
7 0				71.3	71.2	- 0.1
4 0		Oct-Dec		48.0	53.0	+10.4
5 0				52.3	51.6	- 1.4
7 0				51.4	51.8	+ 0.8

* 0=forest/Idle 1=agricultural cropland 2=agricultural pasture 3=urban residential 4=urban employment

** % Difference = $\frac{\text{LANDSAT-NVPDC}}{\text{NVPDC}}$

Section 6.0 Summary and Analysis of Simulation Results

The simulation results fall naturally into those relating to the hydrologic response and those relating to the production and transport of water quality constituents. Each of these is discussed separately below.

Section 6.1 Hydrologic results

The changes in hydrologic response were caused by changes in area assigned to the subbasin and in the changes in the overland flow slope and interflow parameter in subbasin 4. The changes in response for subbasin 4 are summarized in Tables 6.1 and 6.2. These show the per unit area changes in flow components and changes in the total outflow volume. The parameter differences in subbasin 4 changed the distribution of flow among the possible flow paths. The reduction in overland flow volume for the LANDSAT case would lead to a reduction in the carrying capacity for sediment. The increase in overland flow slope, however, may cause higher velocities which would compensate this to some extent. The total per unit area production of water was essentially unchanged by these changes in parameters. The total volume comparison for subbasin 4 also shows that both sets of parameters perform equally well when compared to the observed streamflow from this subbasin.

Section 6.2 Water Quality Results

A summary comparison of the results from the two parameter sets will be presented first followed by a detailed examination of some of the more striking differences revealed by the summary comparison.

Section 6.2.1 Summary Comparison

Table 6.3 shows the outflow of water and constituent as simulated at reach 65. The differences are all small except for changes in the quantity of orthophosphates (PO₄) and biological oxygen demand (BOD). Table 6.4 shows the annual production of constituent from each of the subbasins. The changes are small for the most part, but several exceptions exist. Most notable is the large increase in sediment production in subbasin 5 without a corresponding change in the production of BOD and several other constituents. These and other effects are discussed in detail below.

TABLE 6.1 ANNUAL PER UNIT AREA RESPONSE OF PERVIOUS LAND SURFACE
HYDROLOGY IN SUBBASIN 4

Water Year	Flow Component (inches)					
	Overland Flow		Interflow		Baseflow	
	NVPDC	LANDSAT	NVPDC	LANDSAT	NVPDC	LANDSAT
1975	4.82	4.37	9.25	9.87	6.29	6.17
1976	2.41	2.22	4.73	5.04	5.50	5.40
					20.36	20.42
					12.64	12.66

Table 6.2 COMPARISON OF MONTHLY OUTFLOW VOLUME FOR SUBBASIN 4, IN INCHES

Water Year	Data Source	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	TOTAL
1975	NVPDC	0.06	0.07	4.13	1.38	1.48	2.98	0.86	0.68	0.94	1.44	1.20	5.36	20.58
	LANDSAT	0.08	0.10	4.23	1.40	1.49	3.03	0.86	0.70	1.01	1.50	1.23	5.47	21.12
	Observed*	0.29	0.29	4.14	1.73	1.47	3.24	1.22	1.14	0.49	1.05	1.07	5.43	21.55
1976	NVPDC	1.76	0.66	0.77	5.12	1.44	0.173	1.33	0.30	0.25	0.26	0.09	0.16	12.90
	LANDSAT	1.79	0.67	0.80	5.16	1.44	0.75	1.34	0.33	0.29	0.30	0.11	0.22	13.21
	Observed*	2.00	1.22	1.27	3.91	1.47	1.19	1.83	0.69	0.52	0.59	0.16	0.31	15.18

*Data from USGS stream gage 01656500, Broad Run at Buckland, VA.

TABLE 6.3 COMPARISON OF CONSTITUENT OUTFLOW AT THE END
OF REACH 65

Constituent	Water	NVPDC	LANDSAT	% Difference(*)
Flow	75	106,000	108,000	+2
(Ac-ft)	76	66,600	67,700	+2
BOD	75	459	411	-10
(Tons)	76	299	265	-11
DO	75	1,470	1,500	+2
(Tons)	76	1,044	1,060	+2
NH3	75	13.20	12.50	-5
(Tons)	76	8.45	7.95	-6
NO3	75	104	100	-4
(Tons)	76	55.00	52.50	-5
Org-N	75	70.40	71.50	+2
(Tons)	76	40.50	40.20	-1
Org-P	75	7.77	7.35	-5
(Tons)	76	4.22	3.99	-5
PO4	75	15.00	13.70	-9
(Tons)	76	7.48	6.67	-11

*Defined as $\frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}}$

NVPDC

TABLE 6.4
COMPARISON OF CONSTITUENT LOAD FROM EACH SUB-BASIN

Constituent	Water Year	Sub-Basin 4		Sub-Basin 5		Sub-Basin 7*	
		NVPDC	LANDSAT	NVPDC	LANDSAT	NVPDC	LANDSAT
Flow (Ac-ft)	75	54380	57700	30100	28900	20600	21000
	76	33600	35700	16600	15900	15000	15200
BOD (Tons)	75	639	592	344	303	239	212
	76	285	270	191	163	193	172
DO (Tons)	75	640	670	372	358	253	257
	76	422	438	221	212	192	196
NH ₃ (Tons)	75	12.0	11.1	4.4	4.8	3.5	3.4
	76	6.0	5.6	2.6	2.6	2.7	2.7
NO ₃ (Tons)	75	37.1	32.5	13.3	15.1	11.1	10.7
	76	19.5	18.3	7.7	8.1	8.5	8.2
Org-N (Tons)	75	118.0	115.0	51.5	53.1	38.2	36.9
	76	55.9	54.0	26.3	26.1	29.7	29.0
Org-P (Tons)	75	12.5	15.9	5.1	6.1	4.3	4.0
	76	5.4	4.8	2.6	2.8	3.3	3.1
PO ₄ (Tons)	75	6.1	3.9	1.4	2.3	1.5	1.4
	76	2.6	1.8	0.7	1.0	1.1	1.1
SED (Tons)	75	15300	13700	5100	7920	4390	4360
	76	5730	4760	2150	3200	3560	3640

* Load corrected to exclude Sub-basin 7 area not contributing to Reach 65.

Section 6.2.2 Detailed Comparison

Land surface response :

The effects of differing land use distribution in the two databases are often complex. A more detailed analysis of the annual loading of sediment, biological oxygen demand (BOD), ammonia (NH₃) and orthophosphates (PO₄) from the land surfaces in water year 1975 is given for illustration.

The input for HSPF modelling based on remotely sensed information contained several changes in the parameters related to the above loadings when compared with the input from the conventionally obtained data alone. These changes are detailed in Table 5.13. Parameters were modified in the forest/idle land segment in all three subbasins, and in the urban residential segment in subbasins 4 and 5. The resultant changes in per unit area and per land segment response are shown in Table 6.5. The loading rate and potency parameter changes affected the production of sediment associated constituents only. The per unit area dissolved constituent loads from subsurface flows (groundwater and interflow) and from impervious surfaces were not changed. Table 6.6 shows the approximate percentage of constituent loading attributable to sediment associated constituents.

Because of the extensive changes in subbasin 4 in both water quality and hydrology parameters, the difference in simulation results via the two databases tends to be relatively large. In contrast, there were minimal parameter changes in subbasin 7. These included the parameters of sediment loading and the BOD potency factor for the forest/idle area, and neither change was in excess of five per cent (Table 5.13). Results for this subbasin show that, with the exception of BOD, there were only very small differences between the simulations.

Table 6.7 gives the details of BOD production within subbasin 7. With the exception of the pervious forest/idle land, all annual per unit area production values are identical. The change in the forest/idle land use resulted in an increase of roughly 5% of the total in the per segment production of BOD when the remotely sensed database was used. The total land area of the subbasin also increased by 3%. Yet overall BOD load per subbasin actually decreased. This shows the effects of changing land use distribution within a subbasin even with almost identical parameters.

On the other hand, Tables 6.8 and 6.9 show a 55% increase in annual sediment load in subbasin 5. Yet there is no corresponding changes in the loading of other quality constituents. Instead, the changes here varied from -12% to +63%. These differences are shown in detail in Table 6.9, and discussed separately below.

1. Sediment

The total subbasin sediment output increased by 55% in using the remotely sensed database. Most of this increase is contributed by the pervious portion of the agricultural cropland. In this segment of land use, there was a threefold increase in acreage although the per unit area production

Table 6.5 ANALYSIS OF ANNUAL LOAD OF SEDIMENT AND SELECTED CONSTITUENTS BY SUBBASINS, WATER YEAR 1975**

Constituent	Sub-Basin	NVPDC			LANDSAT			Diff. (%) ^a
		Unit		Total Production (tons)	Unit		Total Production (tons)	
		Production (tons/acre)	Area (acre)		Production (tons/acre)	Area (acre)		
SEDIMENT	4	0.48	32,136	15,268	0.42	32,678	13,729	-10.0
	5	0.34	14,933	5,117	0.55	14,473	7,923	+5.5
	7	0.25	13,631	4,385	0.32	13,990	4,460	+1.7
BOD	4	40.0	32,136	642	36.4	32,678	595	-7.3
	5	46.1	14,932	344	41.9	14,473	303	-11.9
	7	35.2	13,631	239	30.4	13,990	212	-11.3
NH ₃	4	0.75	32,136	12.1	0.68	32,678	11.1	-8.3
	5	0.59	14,932	4.4	0.66	14,473	4.8	+9.1
	7	0.50	13,631	3.4	0.49	13,990	3.2	-5.9
PO ₄	4	0.38	32,136	6.1	0.24	32,678	3.9	-36.1
	5	0.19	14,932	1.4	0.32	14,473	2.3	+64.3
	7	0.23	13,631	1.5	0.20	13,990	1.4	-6.7

$$* \text{ Defined as } \frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}}$$

★★ Load corrected to exclude subbasin 7 area not contributing to channel reaches in study.

Table 6.6 MEAN PERCENTAGE OF TOTAL WATER QUALITY CONSTITUENT ASSOCIATED WITH SEDIMENT LOAD FROM PERVIOUS LAND AREAS. AGGREGATED FOR ALL SUBBASINS IN BOTH DATABASES, WATER YEAR 1975

LAND USE	BOD		NH ₃		PO ₄	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
0 Forest	67	4.5	30	40	12	1.5
1 Agricultural Cropland	84	4.6	80	3.5	96	0.6
2 Agricultural pasture	86	3.5	41	6.7	55	7.5
3 Urban Residential	79	4.9	52	7.2	70	6.4
4 Urban Employment	70	5.3	41	6.4	58	6.7
Total (pervious & impervious)*	66	6.0	43	5.5	72	11.0

* By definition, the impervious areas only contribute dissolved constituents, although pervious areas contribute both dissolved and sediment associated constituents.

Table 6.7 DETAILS OF ANNUAL BOD LOAD FROM LAND USE SEGMENTS IN SUBBASIN 7 FOR WATER YEAR 1975 ***

Land Use	Unit Production** (tons/acre)	NVPDC		LANDSAT		Difference * (%)
		Area (acres)	Total Production (tons)	Area (acres)	Total Production (tons)	
0 P	17.40(17.37)	6,967	60.6	8,775	76.2	+26
0 I	25.22	78	0.9	98	1.2	+33
1 P	35.49	1,733	30.8	1,617	28.7	- 7
1 I	25.22	19	0.3	18	0.2	-33
2 P	49.97	3,394	84.8	2,169	54.2	-36
2 I	25.22	38	0.4	24	0.3	-25
3 P	33.16	490	8.0	528	8.7	+ 9
3 I	56.04	111	3.0	194	5.4	+80
4 P	21.66	268	2.9	170	1.8	-38
4 I	178.70	533	47.7	398	35.5	-26
Total	35.2(30.4)	13,631	239.4	13,990	212.2	-11

P = Pervious land I = Impervious land

* Defined as $\frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}}$

** Where it differs, the value for the LANDSAT database is shown in parenthesis.

*** Load corrected to exclude Subbasin 7 area not contributing to channel reaches in study.

Table 6.8 COMPARISON OF ANNUAL SEDIMENT AND CONSTITUENT
LOAD FROM SUBBASIN 5, WATER YEAR 1975

Constituent (tons)	NVPDC	LANDSAT	% Difference *
Sediment**	5,100	7,920	+ 55
DO	372	358	- 4
BOD**	344	303	- 12
NH ₃ **	4.42	4.78	+ 8
NO ₃	13.3	15.1	+ 14
ORG-N	51.5	53.1	+ 3
PO ₄ **	1.41	2.30	+ 63
ORG-P	5.08	6.08	+ 20

* Defined as $\frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}}$

** Analyzed in detail.

Table 6.9 DETAILED COMPARISON OF ANNUAL LOAD OF SEDIMENT AND SELECTED CONSTITUTENTS FROM SUBBASIN 5, WATER YEAR 1975.

a. Sediment

Land Use	NVPDC			LANDSAT		
	Unit Production (tons/acre)	Area (acre)	Total Production (tons)	Unit Production (tons/acre)	Area (acre)	Total Production (tons)
0 P	.2973	8,139	2420	.3050	9,099	2,775
I	.05615	91	5	.05615		5
1 P	3.005	594	1785	3.005	1,497	4,498
I	.05615	7	0.4	.05615	17	1
2 P	.1195	4,285	512	.1195	2,144	256
I	.05615	48	3	.05615	24	1
3 P	.1227	318	39	.1260	508	64
I	.6147	43	26	.6147	179	110
4 P	.1306	513	67	.1036	271	28
I	.2901	894	259	.2901	632	183
TOTAL P	0.348	13,849	4823	0.564	13,519	7,622
TOTAL I	0.271	1,083	294	0.316	953	301
TOTAL	0.343	14,932	5117.	0.547	14,473	7,923

Annual Total $\frac{\text{LANDSAT}-\text{NVPDC}}{\text{LANDSAT}} = +55\%$

0 = forest/idle 1 = agricultural cropland 2 = agricultural pasture
 3 = urban residential 4 = urban employment
 P = pervious area I = impervious area

Table 6.9 (Continued)

b. BOD

Land Use	NVPDC			LANDSAT		
	Unit Production (lbs/acre)	Area (acre)	Total Production (lbs)	Unit Production (lbs/acre)	Area (acre)	Total Production (lbs)
0 P	24.7	8,139	200,626	29.9	9,099	226,838
I	25.7	91	2,334	25.7	101	2,591
1 P	66.0	594	39,222	66.0	1,497	98,847
I	25.7	7	180	25.7	17	436
2 P	62.3	4,285	267,041	62.3	2,144	133,614
I	25.7	48	1,231	25.7	24	616
3 P	39.3	318	12,482	40.1	508	20,376
I	59.6	43	2,563	39.6	179	10,670
4 P	28.1	513	14,405	28.1	271	7,610
I	166.3	894	148,672	166.3	632	105,101
TOTAL P	38.5	13,849	533,776	36.0	13,519	487,285
I	143.1	1,083	154,980	125.3	953	119,414
TOTAL	46.1	14,932	688,756	41.9	14,473	606,699

Annual Total $\frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}} = -12\%$

c. NH_3

Land Use	NVPDC			LANDSAT		
	Unit Production (lbs/acre)	Area (acre)	Total Production (lbs)	Unit Production (lbs/acre)	Area (acre)	Total Production (lbs)
0 P	.4398	8,139	3,580	.4433	9,099	4,034
I	.2855	91	26	.2855	101	29
1 P	1.773	594	1,053	1.773	1,497	2,654
I	.2855	7	2	.2855	17	5
2 P	.5332	4,285	2,285	.5322	2,144	1,143
I	.2855	48	14	.2855	24	7
3 P	.6622	318	211	.6722	508	341
I	1.711	43	74	1.711	179	306
4 P	.5429	513	279	.5429	271	147
I	1.490	894	1,332	1.490	632	942
TOTAL P	.53	13,849	7,407	.62	13,519	8,320
TOTAL I	1.34	1,083	1,447	1.35	953	1,288
TOTAL	.59	14,932	8,854	.66	14,473	9,608

Annual Total $\frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}} = +8\%$

0 = forest/idle 1 = agricultural cropland 2 = agricultural pasture
 3 = urban residential 4 = urban employment
 P = pervious area I = impervious area

Table 6.9 (Concluded)

d. PO₄

Land Use	NVPDC			LANDSAT		
	Unit Production (lbs/acre)	Area (acre)	Total Production (lbs)	Unit Production (lbs/acre)	Area (acre)	Total Production (lbs)
0 P	.04131	8,139	336	.04147	9,099	377
I	0	91	0	0	101	0
1 P	2.250	594	1,337	2.250	1,497	3,368
I	0	7	0	0	17	0
2 P	.1450	4,283	621	.1450	2,144	311
I	0	48	0	0	24	0
3 P	.3335	318	106	.3401	508	173
I	.8547	43	37	.8547	179	153
4 P	.2333	513	120	.2333	271	63
I	.2975	894	266	.2975	632	188
TOTAL P	.18	13,849	2,520	.32	13,519	4,292
TOTAL I	.28	1,083	303	.36	953	341
TOTAL	.19	14,932	2,823	.32	14,473	4,633

Annual Total $\frac{\text{LANDSAT} - \text{NVPDC}}{\text{NVPDC}} = +63\%$

0 = forest/idle 1 = agricultural cropland
 3 = urban residential
 P = pervious area

2 = agricultural pasture
 4 = urban employment
 I = impervious area

remained constant. The impact of this is lessened by the other significant contribution of sediment from forest/idle land. The latter load is of comparable magnitude, however, it represents only a 15% increase over the load from the simulation with the NVPDC parameters.

2. BOD

Rather surprisingly, the BOD load actually decreased by 12% in using the parameter set based on remotely sensed data. A look at the load breakdown per land use segment shows that two major areas of BOD contribution decreased due to acreage changes alone. This resulted from a halving of the pervious agricultural pasture acreage, where the unit area load is second highest in the subbasin (about 60 lbs/acre), and a 30% decrease in the impervious urban employment area, where the annual BOD load is about 170 lbs/acre. The extra acres were distributed mainly to the pervious agricultural cropland and forest/idle classes. In the latter, the increase in acreage was compounded by a slight increase in per unit area load. However, these effects were not enough to offset the decrease mentioned above.

3. NH3

The annual load of this constituent increased by 8% in the simulation with the remotely sensed database. The major increase in load due to increasing agricultural cropland area is balanced by a similar decrease in the agricultural pasture (segment 52) area.

4. PO4

The increase in the PO4 load is closest to the change in sediment load. The increase in acreage of agricultural cropland with its high per unit area load (at 2.25 lbs/acre, the highest by almost an order of magnitude) accounts for much of the change.

Thus in summary, while parameter and per unit area production changes for a given constituent for a land use may be small or even non-existent, the acreage of the land use and the magnitude of the per unit area production affect the total subbasin output in various ways.

Stream response :

The simulations produced output time series of hydrologic and water quality conditions at the outflow of Reach 65, i.e., at the sampling point, over the water years 1975 - 1976. Time series showing flow rate, water temperature and the concentrations of dissolved oxygen (DO), biological oxygen demand (BOD), ammonia (NH3) and orthophosphate (PO4) are shown in Figures 6.1 to 6.12. Each figure shows three time series : the baseline simulation using the conventional database, the simulation with remotely sensed land use information, and the actual values observed at the sampling point by the NVPDC. The symbols used for these time series are given in Table 6.10.

Since only minor changes were made to hydrologic parameters in subbasin 4, the overall flow rate and water temperature for both databases are,

predictably, very similar. Both sets of time series compare favourably with the observed data.

Inspection of the DO time series show that these values are almost identical with each other and very close to saturation levels. They also show a clear inverse relationship to the water temperature graphs, reflecting the natural relationship between water temperature and saturation levels of dissolved oxygen. Therefore, simulated DO values also compare favourably with the observed values.

Time series plots of BOD, NH₃ and PO₄ also did not show significant differences between simulations based on the two databases. Simulated and observed results frequently differ. The observed data available for comparative purposes here is severely limited to selected runoff events, so that the frequency of data is on the order of two to four days per calendar month. Comparisons with intermittent data is always difficult since differences in timing of events may be confused with the magnitude of events. In addition, the low levels of correspondence of the simulated to the observed may be attributed to the low concentrations of these constituents. These concentrations are at the lower end of the resolution of the standard methods used, where the precision of analyses as measured by one relative standard deviation is usually 20% or worse.

The consistent difference between the simulated responses for PO₄ and NH₃ which appeared after the first few months of the simulation and disappeared near the end of the simulation can be related to the meteorologic conditions in the watershed during the simulation period. The lines agree during periods of extended low flows and diverge during periods of normal high and low flow. The first few months and the last few months of the selected simulation period contained only low flows. The remainder of the simulation period contained significant high flows which maintained the small difference in response between the NVPDC and the LANDSAT parameter cases.

In brief, the simulations based on the conventional and remotely sensed data sets performed equally. However, though differences between the two databases produced many localized changes in land use distribution and water quality parameters, the net effect was low as these tended to cancel each other. Major factors influencing this include the unit area response the acreage of each land use segment, and the effects of components not affected by land use, such as dissolved quality constituents.

Note: After acceptance of the final report, confirmation of a revised version of the observed orthophosphate-P concentrations for Reach 65 were received. Figures 6.11 and 6.12 were revised, but the conclusion that the conventional and remotely sensed databases performed equally poorly in orthophosphate simulation when compared with the observed values remains unchanged.

It should also be noted that HSPF uses a one layer representation of Lake Manassas (Reach 65). Significantly better correspondence to observed values was obtained during simulation with the conventional database in NVPDC

TABLE 6.10 LIST OF SYMBOLS USED IN FIGURES 6.1 TO 6.12

- Simulation with conventional parameter set
- ▲ Simulation with parameter set partially obtained through remote sensing
- + Observed data

studies using HSPX, a proprietary model developed by Hydrocomp. HSPX is similar to HSPF but has a two layer lake representation. By using the same watershed model, in this case HSPF, for both databases, possible bias in favor of either mode of data collection is eliminated.

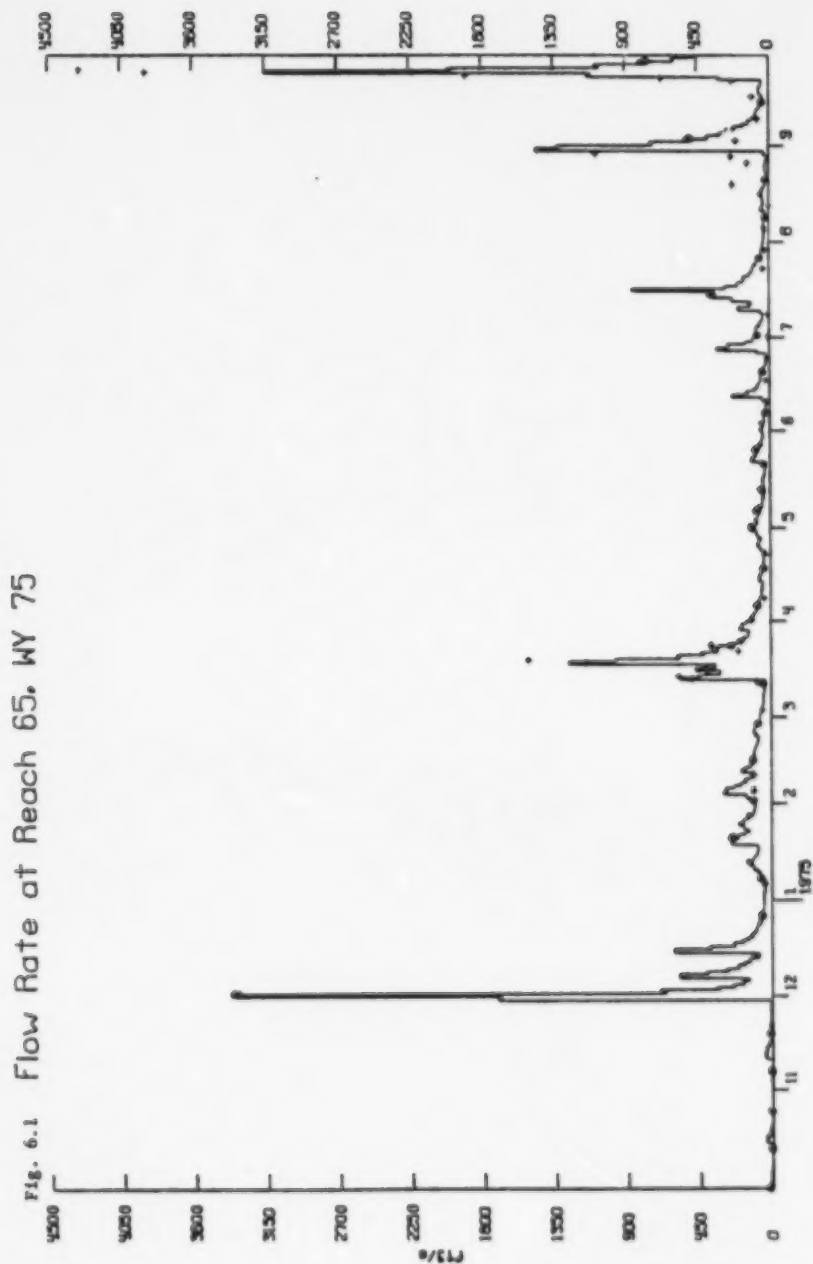
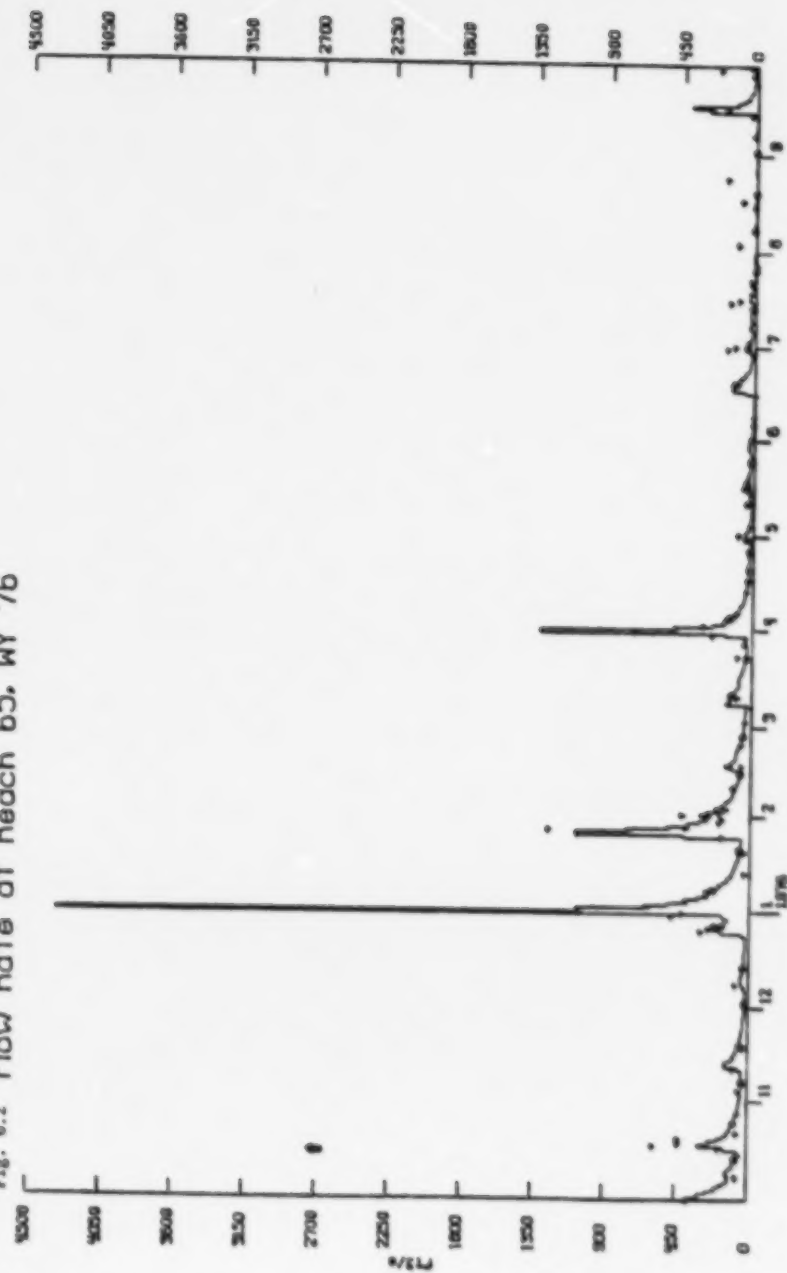


Fig. 6.1 Flow Rate at Reach 65, WY 75

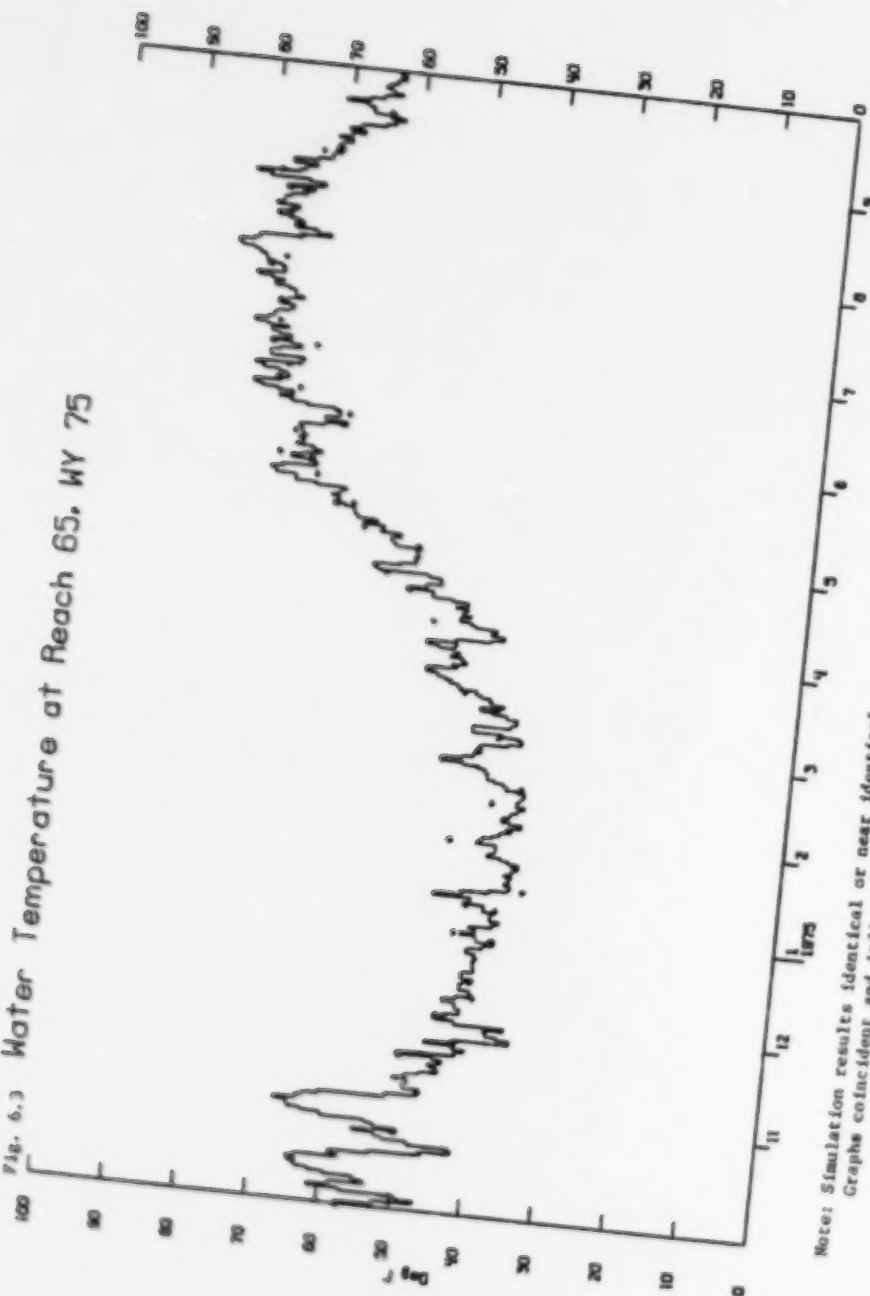
Note: Simulation results identical or near identical.
Graphs coincident and indistinguishable.

Fig. 6.2 Flow Rate at Reach 65. WY 76



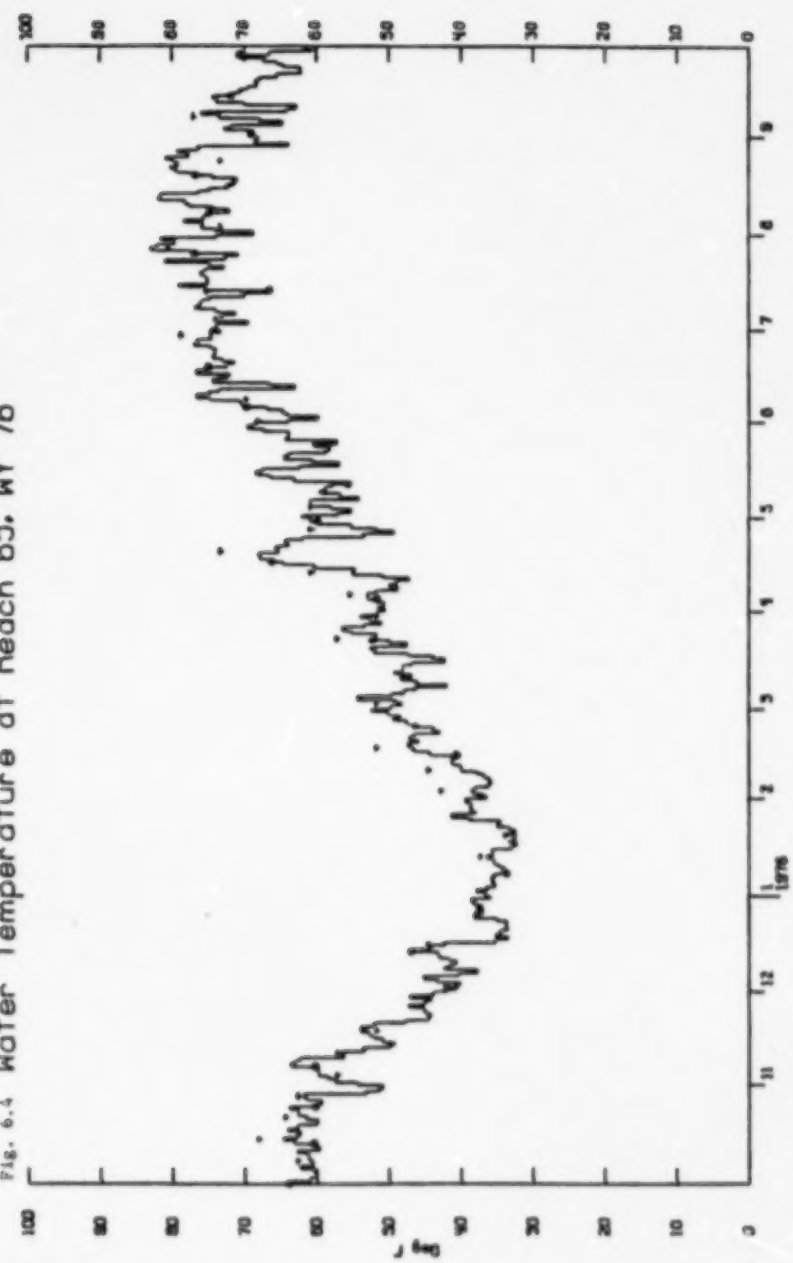
Note: Simulation results identical or near identical.
Graphs coincident and indistinguishable.

Fig. 6.3 Water Temperature at Reach 65. WY 75



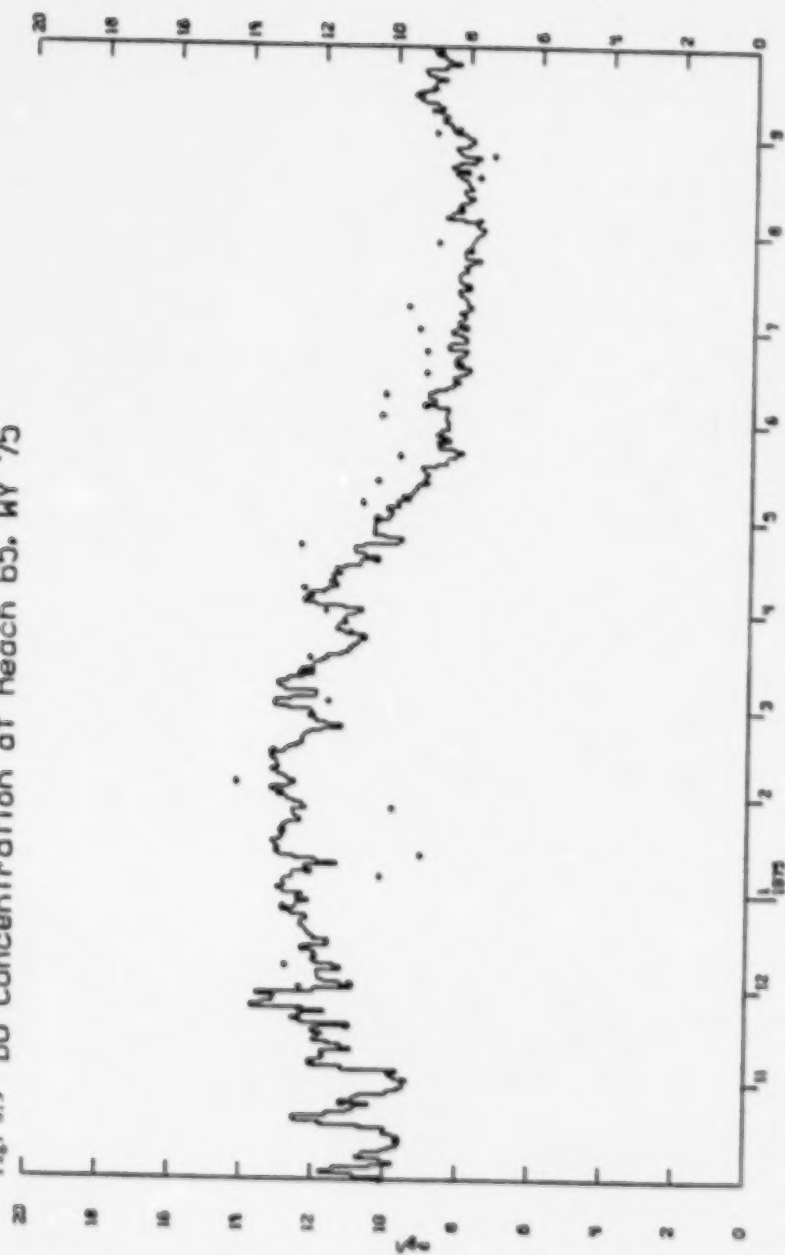
Note: Simulation results identical or near identical.
Graphs coincident and indistinguishable.

Fig. 6.4 Water Temperature at Reach 65, WY 76



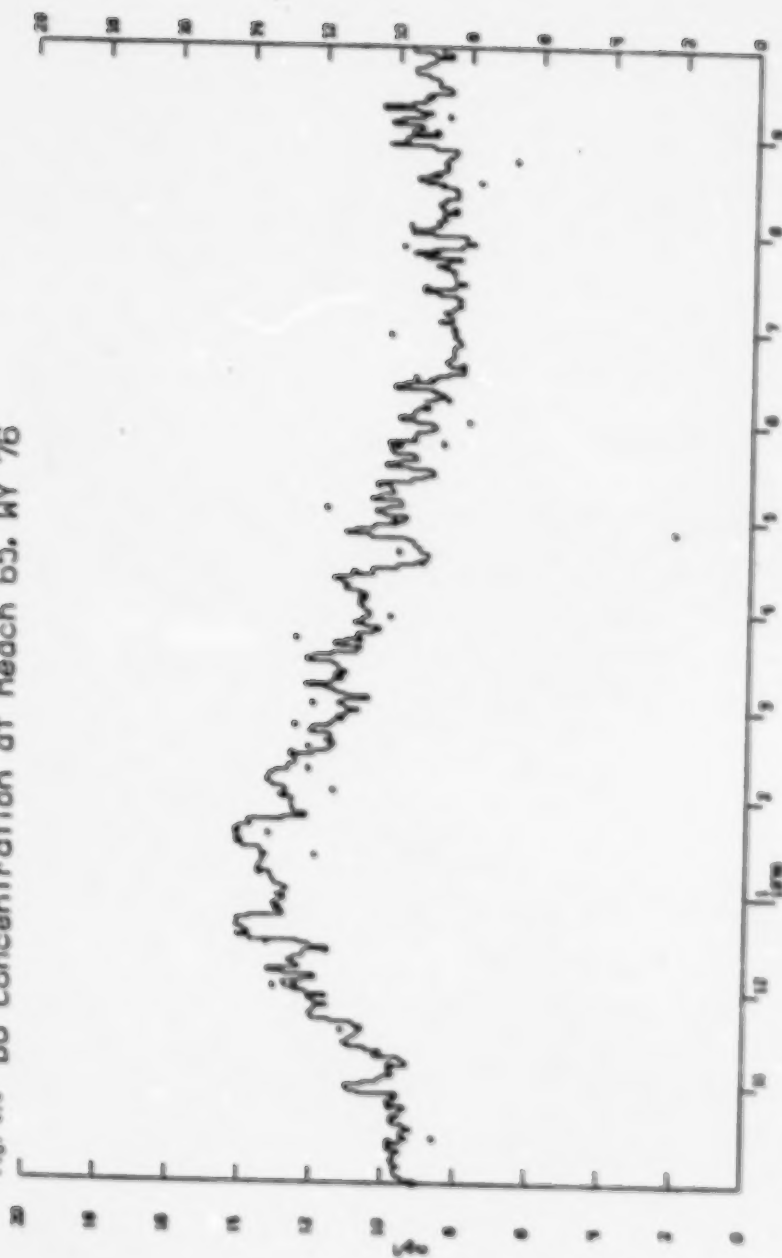
Note: Simulation results identical or near identical.
Graphs coincident and indistinguishable.

Fig. 6.5 DO Concentration at Reach 65. WY 75



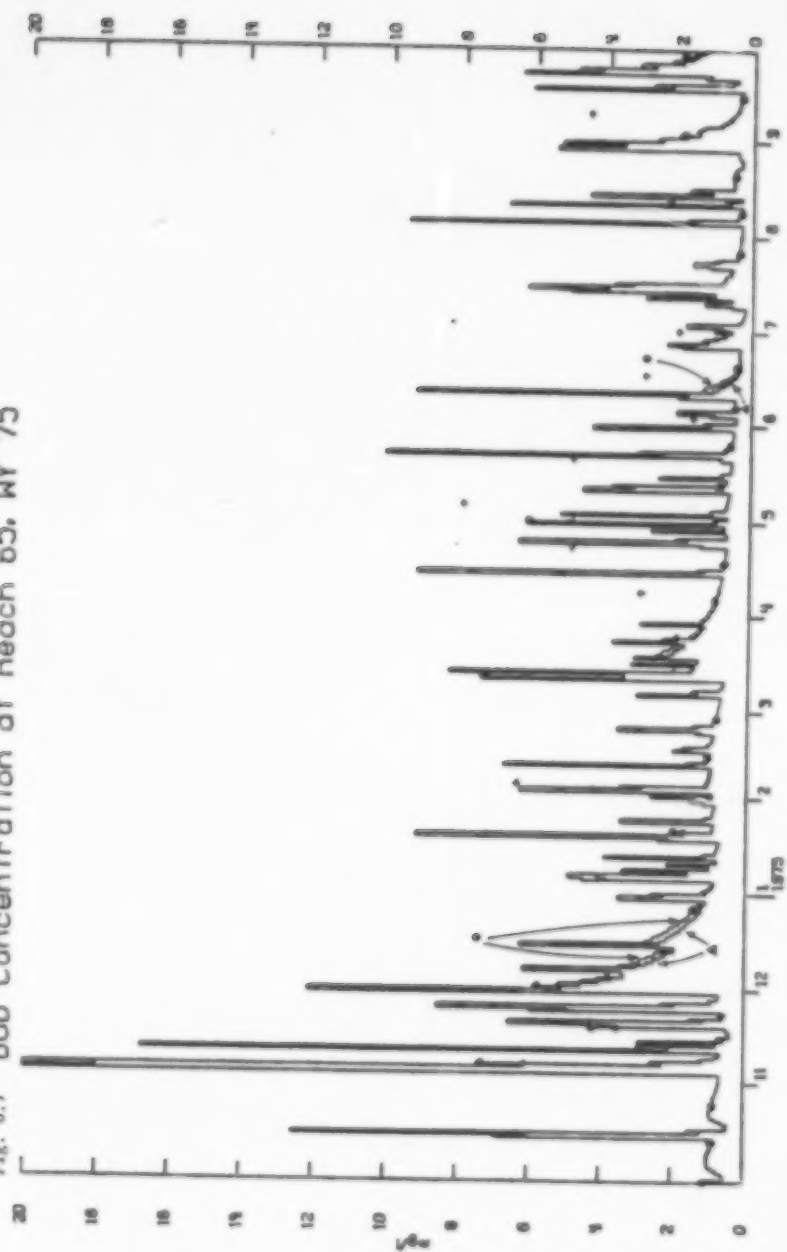
Note: Simulation results identical or near identical.
Graphs coincident and indistinguishable.

Fig. 6.6 DO Concentration at Reach 65, HY 76



Note: Simulation results identical or near identical.
Graph column idiom and indistinguishable.

Fig. 6.7 BOD Concentration at Reach 65, WY 75



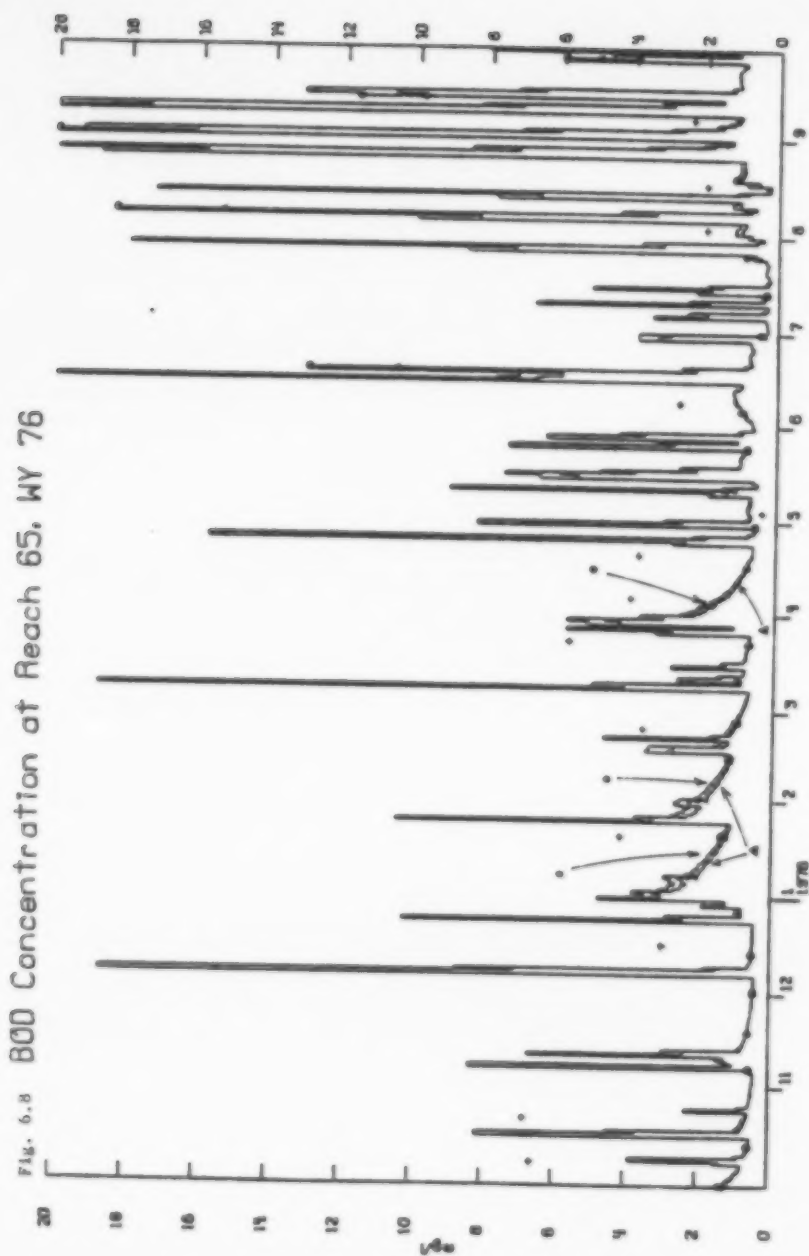


Fig. 6.9 NH₃ Concentration at Reach 65, WY 75

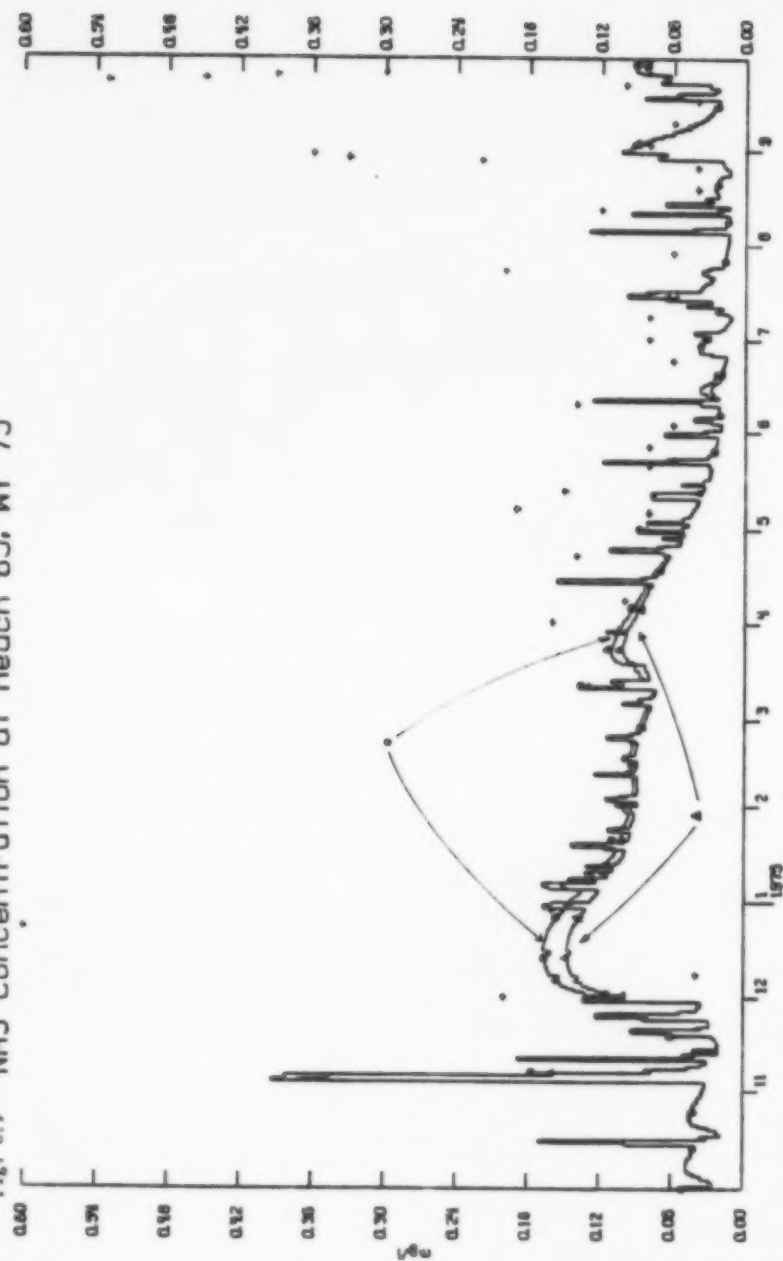


Fig. 6.10 NH₃ Concentration at Reach 65, WY 76

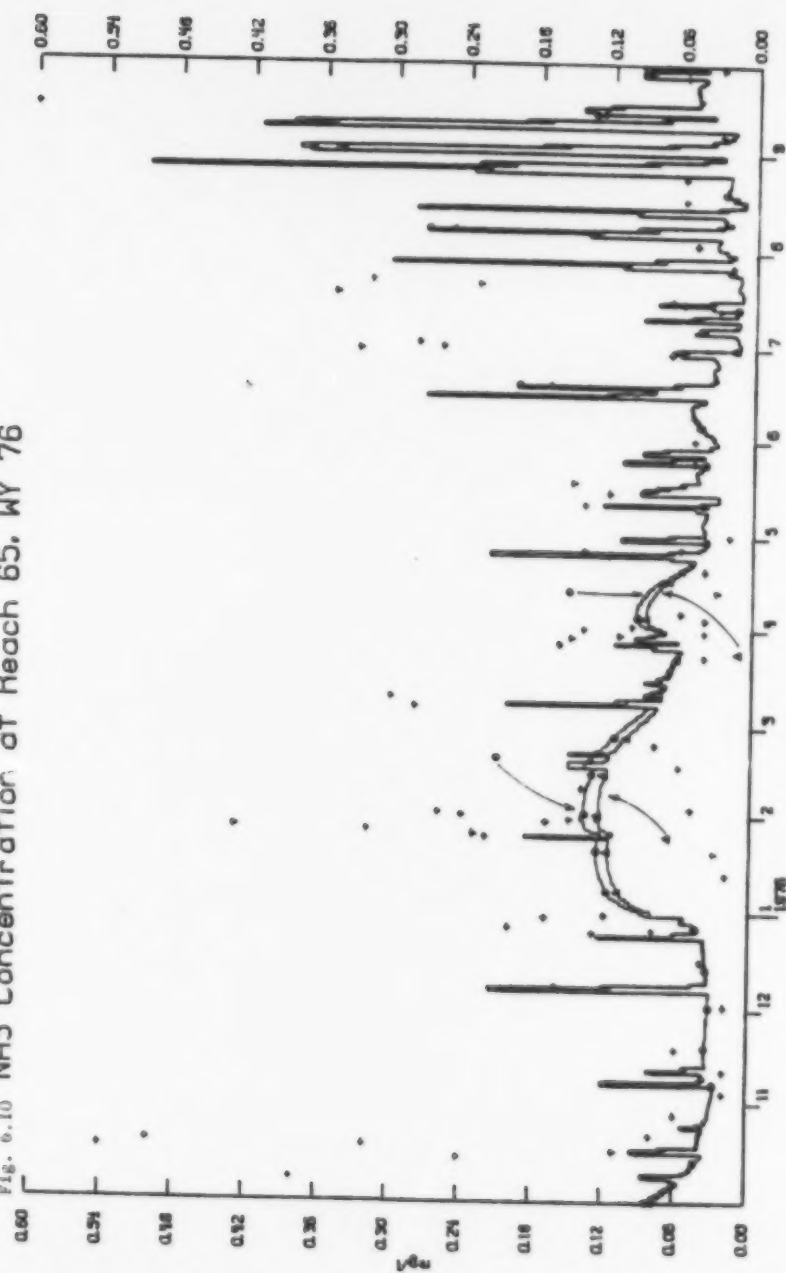


Figure 6.12 PO₄ Concentration at Reach 65, WY 76

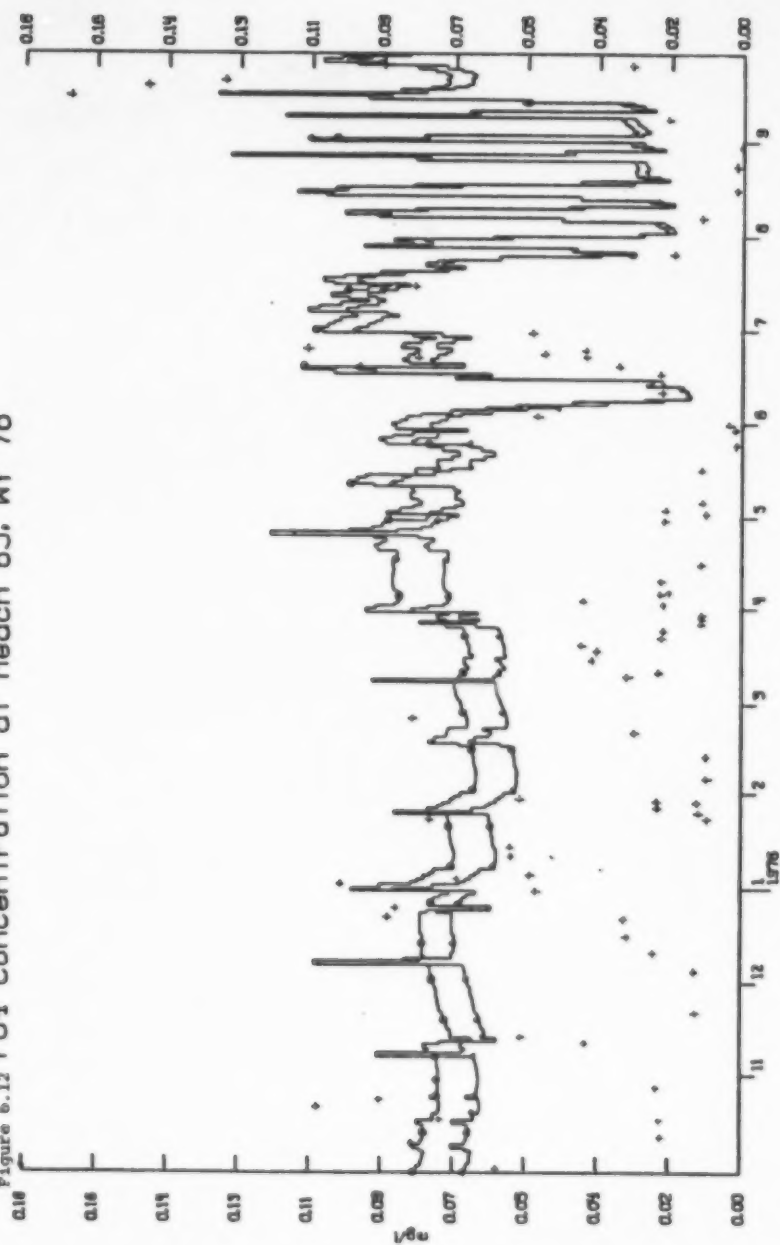
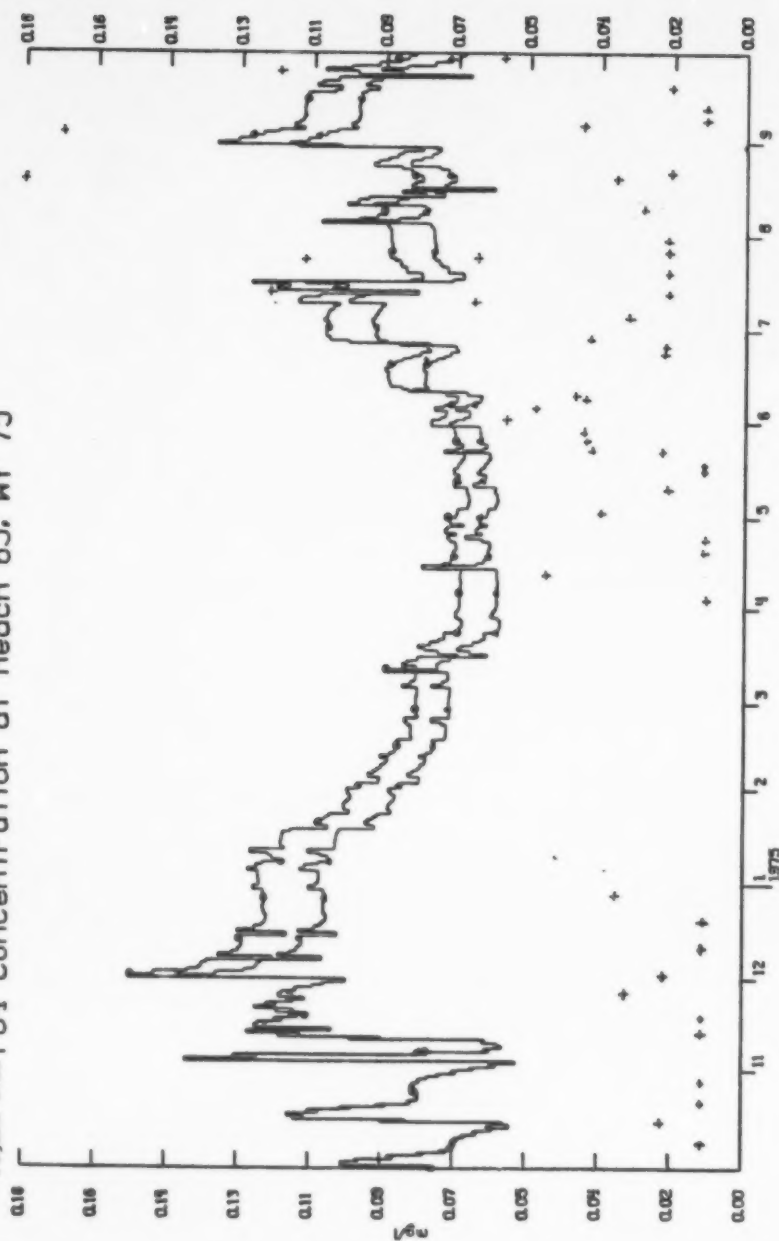


Figure 6.11 P04 Concentration at Reach 65, WY 75



Section 7 Cost effectiveness of current and other studies

Cost estimates for setting up the conventional and remotely sensed land use distributions and overlaying these with other terrestrially measured data, were obtained from the NVPDC and NASA respectively. These costs are estimated for the entire Occoquan Basin, not only subbasins 4, 5 and 7, and are shown by tasks in Tables 7.1 and 7.2. From these, labor and computer rates, and other costs were determined (Table 7.3).

Labor estimates were drawn up based on information supplied by these agencies and from the literature. These include certain modifications of the original estimates, and are discussed for each database in detail individually. (See Table 7.4.)

1. Conventional Database

First of all, consultation with other existing agencies (i.e., items IIB and IIE in Table 7.1) were not included since no direct cost was actually incurred. Thus the net labor requirements is 25.5 man-weeks.

Secondly, although a labor mix of 13:12.5 Engineering Intern to Engineer I (or Planner I) in the actual effort by the NVPDC, Communications with the NVPDC indicated that a ratio of 22:3.5 was a practiceable one for an estimate of routine costs. Labor costs were therefore prorated according to this scheme.

The material cost was taken from the mean of the estimated range provided by the NVPDC.

For the compilation of this database, update costs depend largely on the extent of land use changes in the area. Therefore the component of costs is shown as a range of values.

2. Remotely Sensed Database

The original estimate for setting up the remotely sensed database assumed personnel with no experience with IMAGE 100. This was revised to represent an effort with experienced operators unfamiliar with the study area to be consistent with other studies (e.g., Jackson, Ragan and Fitch, 1977). As a result, costs for set up and operations included 3 days of labor plus 2 days of IMAGE 100 computer and operator time. The 8 days of labor for ground truth establishment in the original estimate was not changed.

Cost of updating the database is not dependent on the extent of land use change since the LANDSAT tapes would include the entire basin. Slight modifications of the original update cost estimates were made to reflect personnel experience consistent with the above.

The cost of using IBIS to interface the remotely sensed land use classification with terrestrially obtained data was not included. The use

Table 7.1 ORIGINAL NVPDC COST ESTIMATES FOR A CONVENTIONAL DATABASE IN THE OCCOQUAN

ACTIVITY	MANPOWER REQUIREMENTS (PERSON WEEKS)				TOTAL COST ^a
	ENGINEER I OR PLANNER I	ENGINEER INTERN	DRAFTSMAN		
I. Assemble information for land use map preparation					
A. Assemble maps and photos	1.0				\$ 750
B. Set up property ownership tables for agricultural land in 4 counties	2.0	2.0	1.0		\$ 2,200
C. Prepare base map of watershed					\$2,000-\$3,000 ^b
					TOTAL \$4,950-\$5,950
II. Construct land use map of watershed					
A. Delineate urban land use with aid of aerial photos and existing land use maps; consultation with local staff	2.0	2.0			\$ 2,200
B. Compile property ownership data from local agricultural management agency files	8.0				\$ 6,000
C. Transfer data to base map	2.0	4.0	4.0		\$ 3,000
D. Prepare zipatone mylar					\$ 3,000
E. Review map by local staff (6 jurisdictions)	2.0				\$ 1,500
					TOTAL \$15,700
III. Planimeter land use data by sub-basin (15 subbasins)	0.5	5.0			TOTAL \$ 2,100
TOTALS	17.5	13.0	5.0		\$22,750-\$23,750

a. all costs are in 1980 dollars

b. includes \$1,000-\$2,000 cost for negatives of USGS quads and mylar development

Table 7.2 ORIGINAL NASA COST ESTIMATES FOR A REMOTELY SENSED
DATABASE IN THE OCCOQUAN

R&D Cost Estimate

Two Landsat Tapes (April and August 1975)	\$ 500
Ground Truth (2 persons for 3 days and transportation)	1,600
Ground Truth Map (1 person for 2 days)	300
Image 100 time for two weeks (250 an hour includes operator's time)	20,000
Image 100 Analysis (1 person for two weeks)	1,500
	\$23,900

Operational Cost Estimate

Two Landsat tapes	\$ 500
Image 100 time for two days	4,000
Image 100 analysis for one day	300
	\$ 4,800

Table 7.3 ESTIMATE OF LABOR RATES AND OTHER COSTS

ITEM	CONVENTIONAL DATABASE	REMOTELY SENSED DATABASE
Labor	\$ 80 - 110/day*	\$ 150/day
Computer	---	\$ 250/hour
Materials	\$ 1,500 USGS negatives	\$ 250/landsat tape
Transport	---	\$ 350/person

* Depends on labor mix.

\$ 70/day for Engineering Intern

\$ 150/day for Engineer I and Planner I

Table 7.4 COMPARISON OF ESTIMATED COSTS, CONVENTIONAL VERSUS REMOTELY SENSED DATABASES

a. Set up and Operations

ITEM	CONVENTIONAL DATABASE	REMOTELY SENSED DATABASE
Labor	\$ 9,400 (127.5 days)	\$ 1,650 (11 days)**
Computer	---	\$ 4,000 (16 hours)
Materials	\$ 1,500	\$ 500 (2 tapes)
Transport	---	\$ 700 (2 persons)
TOTAL	10,900	\$ 6,500

b. Update Costs

ITEM	CONVENTIONAL DATABASE	REMOTELY SENSED DATABASE
Labor		\$ 300 (2 days)
Computer	\$3,000 - 7,000 ***	\$1,500 (6 hours)
Materials		\$ 500 (2 tapes)
Transport		---
TOTAL	\$3,000 - 7,000	\$2,300

* Assumes a labor mix of 22:3.4 Engineering intern to Engineer I.

** 8 days for ground trugh establishment, 3 for IMAGE 100 analysis.

*** Depends upon extent of changes which have occurred.

of IBIS resulted in a database of power and flexibility far beyond that of the conventional one, so that the inclusion of IBIS in the cost comparison would result in the evaluation of two databases which are not comparable in abilities. For completeness, the JPL estimates the cost of IBIS use at \$10,000 for a comparable effort; where there is previous experience with the study area, significant savings can be obtained.

Major items not included in the above computation include the development of NVPDC's transformation program to convert land use to hydrologic and water quality parameters and the development of LANDSAT itself. These and other similar costs cannot be satisfactorily factored into that of database development, and the resources are considered to exist as givens.

Under these conditions, the development of a remotely sensed database has significant cost advantages over that of a conventional one in both set up and operations. According to the current estimates, savings of about 40% in set up and operations, and 20% to 70% in updates may be achieved through the use of a remotely sensed database.

Beside the financial costs, the construction of conventional database is favoured by a number of intangible benefits. Land use data collected by the NVPDC may be used by other planning studies and programs in the area. Another benefit of the effort of collection is the interaction with the local planning staff. This improves their understanding of the use of modelling, and in general enhances the credibility of the modelling effort. On the other hand, there are certain repetitious and tedious tasks in the compilation of the conventional land use distribution, for example the calculation of area by planimetry (item III, Table 7.1). As a result, there are often difficulties in retaining trained staff and maintaining a constant quality of data collection.

Remotely sensed database construction also has certain advantages and disadvantages. In contrast to the conventional database, there is little interaction with local agencies during its compilations. Thus there may be less secondary use of the information obtained due to the presence of a greater credibility gap.

As indicated by the present study, LANDSAT coupled with IBIS may give rapid and complete results, obviating much drudgery in the manual effort. As mentioned above, there is therefore an increase in the reliability of the data collection process. The greater objectivity and repeatability of this process are other inherent advantages of the use of a remotely sensed database. Furthermore, the flexibility of IBIS data handling features in particular would also increase cost efficiency.

Secondly, there is a significant economy of scale associated with the construction of a remotely sensed database. The bulk of cost of the conventional effort was for labor, for surveying and setting up land use maps. This is likely to increase at least linearly with increasing area. The comparable remote sensing cost is the much smaller one of acquiring more LANDSAT tapes. Indeed, the logistics involved in setting up a land based

survey and obtaining LANDSAT tapes are orders of magnitude apart. Though there would be some increase in cost for establishing ground truth and setting up the database through IBIS, this is not likely to be a linear function of area.

In addition to the economies of scale, remote sensing has the undeniable advantage of access in both space and time. The Occoquan is a relatively small and very intensively studied watershed, to which most others are not comparable. The latter are often far larger, or in undeveloped or agricultural areas, so that conventional databases cannot be easily set up for lack of access or resources. Also, in areas undergoing rapid changes such as urban development, a conventional survey may not be able to keep its database up to date. These would not present problems for remote sensing.

There are also many potential developments that could further enhance the scope and feasibility of using remotely sensed databases. These are discussed in detail in the following section (Section 8).

Section 8 Potential for the Use of Remotely Sensed Data in Modelling

As was mentioned in the previous section, remotely sensed data on watershed characteristics can improve modelling in two ways. Firstly, improved modelling means increased model accuracy for a given level of effort. The economic savings from using a remotely sensed database therefore improves modelling by providing higher accuracy per dollar investment. Secondly, there is the immediate and practical utility of remote sensing for large or inaccessible watersheds or watersheds needing rapid information updates.

This chapter is a summary of what has been done to assess the potential of remotely sensed data, and the steps needed to more fully assess this potential in future studies. To discuss this topic some background and definitions are needed. The background of the physical processes in hydrology and water quality is included as Section 8.1. Modeling techniques, and definitions of terms used in modelling, are included in Section 8.2. Also included in this section are the descriptions of conventional sources of data used in modelling. Section 8.3 describes the role of remotely sensed data and IBIS data management in continuous simulation modelling. Finally, 8.4 discusses the methods used to assess the potential of remotely sensed data, including what has been done to date on this project, and what could be done in future projects.

Section 8.1 Physical Processes

To understand the hydrologic regime of a basin, one must follow the dynamics of the physical processes as they respond to the time dependent climatic driving forces. The relationships between climatic variables and physical processes are clearly complex and varied. Figure 8.1 shows the key physical processes that are represented in continuous simulation of hydrology and water quality. Table 8.1 lists the key relationships between the various physical processes that control the watershed regime.

For example, for a modelling study of water quality, a conclusion that could be drawn from the above is that the path followed by water in the hydrologic cycle controls the nonpoint sources of pollutants that enter waterways. The starting point to understand the water quality regime of a basin is therefore understanding the hydrologic regime and the amount and timing of flows that move by surface runoff, interflow and groundwater. From this base, one can then calculate the movement of sediment, and the movement of sediment associated and non-sediment associated pollutants such as pesticides, nutrients, and heavy metals from the land surface through the watershed. It is clear that the key parameters in any modelling of watershed water quality must include:

1. Those that control the paths of flow in the watershed.
2. Those that control the loading rate or deposition rate of materials onto the land surface.
3. Those that control degradation of pollutants on the land surface or in waterways.

Parameters for point load pollution from treatment plants are not included in the above list. This is because as water treatment standards have improved, the impact of point load pollution has decreased.

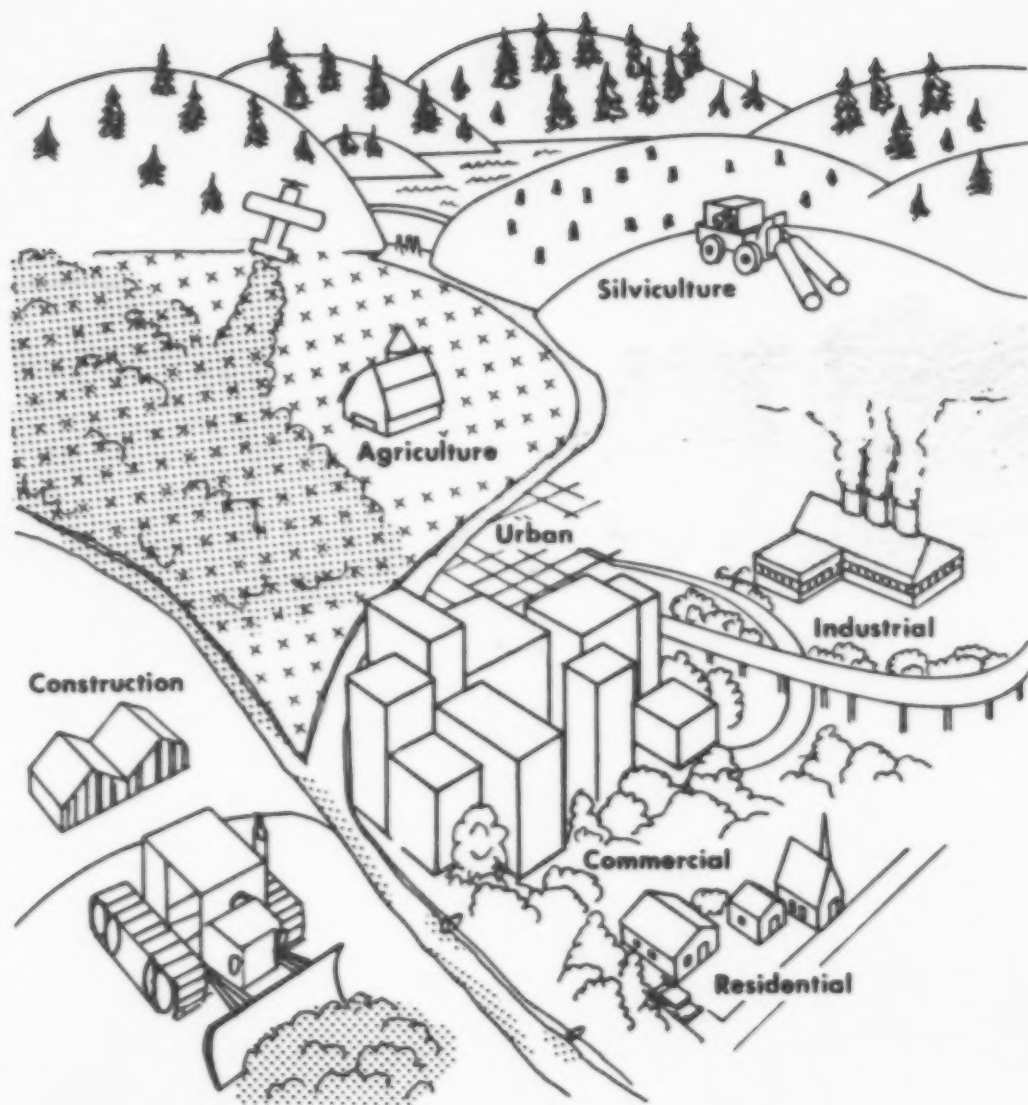


Figure 8.1 Hydrologic and Water Quality Processes.

TABLE 8.1 KEY PHYSICAL PROCESSES IN THE WATERSHED

Process	Frequency	Time Delay	Influenced by	Driving and Dependent Processes
precipitation	<1% of time	--	--	hydrologic cycle and pollutant transport by water or sediment
evaporation	continuous	--	soil moisture storage, vegetation	precipitation
surface runoff	<1% of time at soil surface	<1 hr at surface	soil moisture, soil property, vegetation	precipitation evaporation
groundwater	continuous	2 wk. to years	infiltration vegetation	sediment transport, overland pollutant transport to streams
channel flow	intermittent to continuous	days to weeks	channel shape, slope, instream vegetation	inflow to channels (surface runoff, groundwater)
storage in reservoirs and lakes	continuous	weeks to months	--	storage volume, inflow volume
degradation volatilization	continuous	hours to years	properties of pollutant	environment (moisture, temperature)
solubility	continuous	--	temperature	pollutant transport in solution
detachment/transport	<1% of time with surface runoff	<1 hour	soil properties	precipitation rates, surface runoff
adsorption/desorption	continuous	--	pollutant and soil prop.	moisture movement in soil
air-water interchange	continuous	--	meteorologic conditions	concentration gradients in liquids, gases

Section 8.2 Modelling Techniques

In modelling, the processes shown in Table 8.1 are represented to the best of available knowledge. The model contains in its structure mathematical representations of the key physical processes. For example, overland flow is represented on the land surface using hydrologic equations relating land slope, resistance to flow, and the hydraulic properties of sheet flow. When algorithms for various processes are developed, they are linked together so that the occurrence of overland flow will detach sediment particles on the land surface, causing sediment to move into stream channels. When algorithms for the different physical processes are operated together, deposition of material on the land surface, its degradation by atmospheric processes, its transport with sediment into stream channels, and its transport and degradation in the channel system are all represented and simulated.

Figure 8.2 shows the inputs to and outputs from the modelling process. In this figure, model input is identified as watershed characteristics or time series. This input enters the model structure, and as the model operates, output time series are produced. The output time series are compared with measured field data on water flow, sediment transport, or pollutant concentrations. If the model output and the field measurements do not agree, calibration must take place. In this process, the watershed characteristics that are represented by parameters are modified so that the simulated and observed time series correspond more closely one to another. A change that could be made would be to increase the resistance to overland flow on the land surface to slow down the velocity of surface runoff, and to decrease its volume.

The distinction between input time series such as precipitation or temperature, and watershed characteristics such as land slope is an important one: parameters are calibrated and input time series are not. The parameters represent watershed characteristics, and they are adjusted by calibration when actual watershed characteristics are not easily determined in the field. On the other hand, bias and random error in input time series causes many differences between model outputs and field measurements and these differences cannot be corrected by parameter calibration.

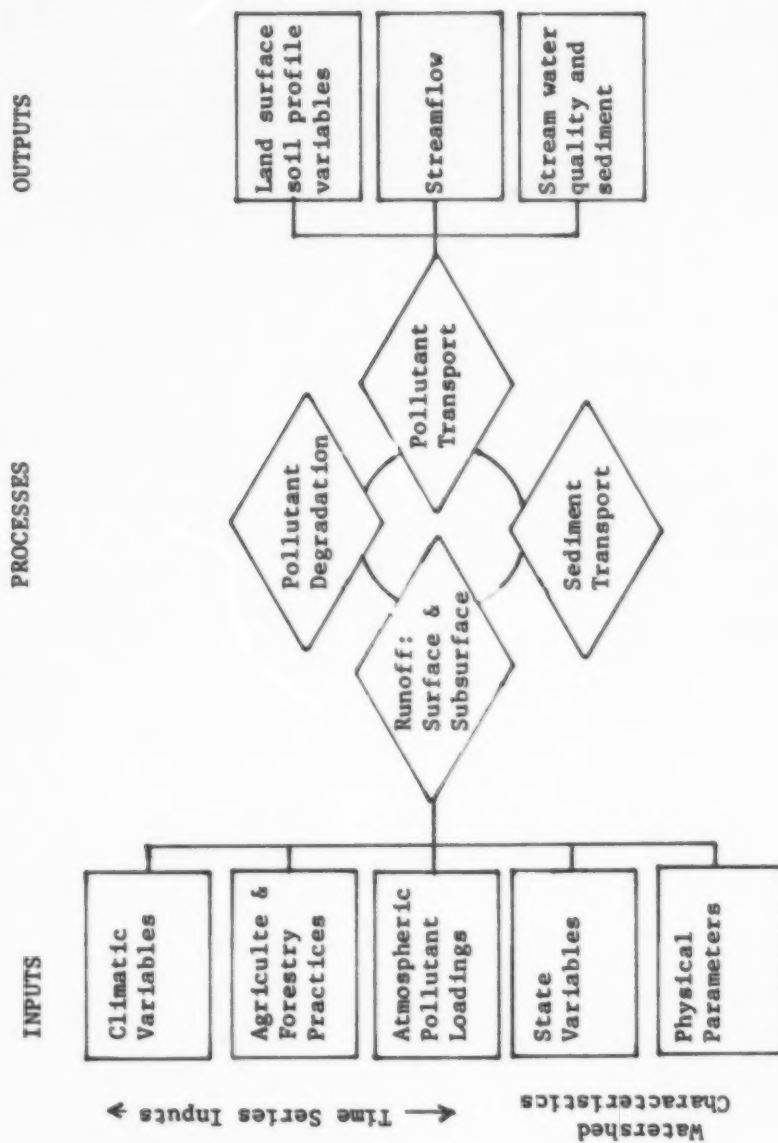


Figure 8.2 Model Inputs, Processes and Outputs.

Section 8.3 The Role of Remotely Sensed Data and IBIS Data Management

Remotely sensed data can be used for parameter derivation or to supply input time series. These applications are discussed in the following subsections.

Section 8.3.1 Parameter Derivation

The parameters used in modelling differ widely in impact: Model response is largely determined by a small subset of the complete parameters list for HSPF. Remote sensing must contribute to these high impact parameters to play an important role in environmental modelling.

The methods by which parameters are derived may be physical or empirical. A physically derived parameter is directly dependent on a measurable physical feature in a watershed. An example is the amount of impervious area in the watershed. An empirically derived parameter may also be physically based, but is not directly related to measurable physical characteristics. Empirically derived parameters depend on relationships between several physical characteristics, or on a combination of physical and climatic characteristics. An example is the interflow parameter, which controls the volume of interflow. A summary of parameter impact and method of derivation for the HSPF system is shown in Fig. 8.3.

Remote sensing can contribute to the definition of both physically derived and empirically derived parameters. Remote sensing is directly useful for physically derived parameters many of the physical characteristics of the watershed can be measured by remote sensing. This type of data is particularly useful in water quality studies where more detailed land use information is required. For example, a major area of environmental concern is pesticide pollution. Since pesticides are often very organism specific, refinement of the remotely sensed agricultural land classification, to give crop type for example, has significant potential in modelling and risk analyses for evaluating pesticide contamination in a given area.

Empirically derived parameters are functions of several physical characteristics. A preprocessor performs a logical compilation of physical characteristics to arrive at a parameter value. The complexity of the preprocessor may vary depending on the nature of the information available and required. A simple example is the conversion of the spectral signatures over an area into the air temperature. The preprocessor used in the current study is relatively complex and indirect. Here, the spectral signatures detected by LANDSAT were converted into land use distributions. These distributions were then used to generate hydrologic and water quality parameters through the use of another, previously developed transformation function or preprocessor.

The derivation of empirical parameters benefit from remotely sensed information where the more complex preprocessors are developed, and where IBIS type data management systems are used to correlate various types of information. IBIS data management is particularly useful because watershed characteristics and parameters vary from point to point on the watershed

surface, so that the empirical parameter must be derived for each unique point by tabulation and transformation of the appropriate geographical data. The integration of calibration techniques, remotely sensed data, IBIS data bases and preprocessor concepts for the high impact, empirically derived parameters offers much promise in the improvement of hydrologic and water quality modelling.

Furthermore, it is possible that generic patterns in correlations between remotely sensed information and required modelling data, independent of specific watersheds, may be established. As has been made obvious, the potential benefits to modelling resulting from this development are very significant.

Section 8.3.2 Improvement of Time Series Information

Remote sensing of input time series for simulation modelling can potentially be used for many variables including the following:

- Precipitation
- Air Temperature
- Actual evapotranspiration
- Timing of agricultural operations
- * Soil moisture
- * Soil surface temperature
- * Snow water equivalent, moisture content and albedo
- * Areal extent of snow cover

The frequency of measurement is important for time series inputs. Typically, measurement frequencies need to be daily or hourly. Variables such as the last four items in this list, marked with asterisks, are accuracy of the modelling effort.

There are other potential applications of remotely sensed information to the construction of the input time series used in modelling. One of the major sources of error in modelling is the reliance on single point measurements to represent large segments of the watershed, since the meteorological and hydrological processes are highly variable from point to point. However, the logistics of setting up and maintaining a large number of measurement stations through large and/or undeveloped watersheds are staggering. As a result, much of the data used in modelling is inferred from a small number of point measurements. In terms of modelling, the measurement of actual geographical variation, say at 1,000 locations in a watershed, with a 20% standard error, is far more useful than measurements at three locations with a 2% standard error.

Continuous monitoring of an area is beyond the capabilities of LANDSAT. However, satellites in synchronous orbit can be used to provide time series of various meteorological and hydrological information such as air temperature, precipitation, radiation and evapotranspiration at many points over these large areas. This development is of especial significance where there are frequent events which are localized, in both time and space. For example, there may be frequent periods of intense rainfall in the hills, but

the rain gauge for the region (on which modelling input relies) is at the airport where rainfall patterns may be quite different. Thus the availability of geographically distributed information can dramatically improve the accuracy of the modelling effort.

In conjunction with the improvement of time series information, another problem associated with complex continuous simulation models is that calculated values for the state variables tend to drift from the observed due to inaccuracies in the time series from which they are computed. Small errors (noise) may accumulate unacceptably over the longer simulation periods. In these cases, due to the random nature of this noise the problem cannot always be solved by parameter adjustment (calibration) or improvements of model structure. This may have serious consequences since many of the simulation algorithms are functions of the state variables; i.e., the accuracy of the simulation is directly dependent on the accuracy of these state variables.

Remotely sensed data can be used to periodically check the simulation by providing updated state variable values. Many of variables would potentially benefit from this, including soil moisture, geographic extent of snow cover, snow depth, soil and water body temperature and tillage and cropping patterns.

Time series inputs from remote sensing may also result in modifications to the model structure. Models are built from algorithms based upon data availability. The availability of previously unavailable time series information could potentially lead to the formulation of new algorithms which more closely and accurately represent watershed processes.

Section 8.4 Methods to Assess the Utility of Remotely Sensed Information and the Use of IBIS Data Management

Techniques to assess the utility of remotely sensed data and IBIS, where "utility" means increased model accuracy or reduced cost, are:

1. To perform a watershed study without remotely sensed data and IBIS, and again with remotely sensed data and IBIS.
2. To perform a study using remotely sensed information only, and a parallel study on a nearby paired watershed with conventional data only.
3. To perform error analyses or sensitivity studies of the effects of physical characteristics and increased accuracy in measuring on model accuracy.

The current study is of the first type.

Furthermore, in the use of remotely sensed information, maximum benefits are associated with better estimates of the high impact parameters. A better value for a low impact parameter might be found, but it will be of more limited utility. Improvements from remotely sensed information are easy to measure in direct physical parameters. Improvements in empirical and calibrated parameters are more difficult because good values for these parameters might be found by calibration, independent of any field information. Figure 8.3 provides a convenient guide relating the method of parameter derivation to the impact of the parameter on water resource systems.

In the current study of the Occoquan, we found clear-cut improvements in the estimation of land use, and therefore also better definition of parameters related directly to land use. The calibrated parameters that depended on preprocessor operations were improved. Proof of accuracy gain based on final results of the study were less clear. This will be the case in any field study, given the current state of field data for water quality.

A possible approach for the future is to perform a study with remotely sensed data only, and then compare the results to nearby watersheds, i.e., a study of the second type. Full scale tests on watersheds give qualitative results which are very helpful to the professional modelling community, even if numerical proof of utility from field measurements of water quality data is not obtained.

Another possible future study is to consider the effects of accuracy gains in the estimation of basic watershed characteristics. The sensitivity of water quality results to changes in watershed characteristics could be investigated; for example, the change in dissolved oxygen that results from a change in land use in the basin.

In summary, for reasons quite independent of remote sensing, it is difficult to show the utility of remote sensing data based on field water quality information. Water quality data is so sparse that a substantial improvement in model accuracy from any source is difficult to demonstrate using field data. However, professional modellers are very comfortable with subjective judgments of utility. Given the many potential benefits to modelling that can be derived from remote sensing, it is very likely that modellers would use remotely sensed inputs extensively based on their subjective judgment that better results or a more accurate reproduction of the prototype watershed was being achieved.

To provide impetus for the use of remotely sensed data for environmental analysis we recommend:

1. More comprehensive field tests, where remotely sensed data is used in modelling analysis, including tests on larger watersheds.
2. Sensitivity studies of the effects of errors in watershed characteristics on model outputs.
3. Research on preprocessor, data base, and calibration techniques for empirical parameters (Figure 8.3).

We believe that remote sensing and IBIS data management will become widely used in environmental modelling over the next decade, and that as routine techniques are developed to use remotely sensed data in modelling, the utility of these data will be clear.

Section 9 References

- Bryant, N.A. and A.L. Zobrist, 1977.
A Geographic Information System Based on Digital Image Processing and Image Raster Data Type. IEEE Transactions on Geoscience Electronics, v GE-15(3):152-159.
- Grimsrud, G.P., D.D. Franz, R.C. Johanson and N.H. Crawford, 1979.
Executive Summary for the Hydrological Simulation Program - Fortran (HSPF). Environmental Research Laboratory, U.S.E.P.A., Athens, GA.
In press, 79 pp.
- Hydrocomp, Inc., 1978.
The Occoquan Basin Computer Model. Calibration, Verification, and User's Manual. Northern Virginia Planning District Commission, Falls Church, VA.
- Jackson, T.J., R.M. Ragan and W.N. Fitch, 1977.
Test of LANDSAT-Based Urban Hydrologic Modeling. J. Water Resources Planning and Management Div., May 1977, pp.141-158.
- Johanson, R.C., J.C. Imhoff and H.H. Davis, Jr., 1980.
User's Manual for the Hydrological Simulation Program - Fortran (HSPF). Environmental Research Laboratory, U.S.E.P.A., Athens, GA.
EPA-600/9-80-015.
- Johanson, R.C. and D.D. Kliewer, 1980.
Maintenance and Testing of Hydrological Simulation Program - Fortran (HSPF). Final Report. Environmental Research Laboratory, U.S.E.P.A., Athens, GA. 83 pp.
- Pabst, A.F. and R.J. Cermak, 1978.
Measurement and Forecasting of Streamflow by Remote Means.
In Proc. of the Military Hydrology Workshop, 17-19 May, 1978, U.S. Army Engineer Waterways Experimental Station, Vicksburg, MS.
- Regional Resources Division, N.V.P.D.C., 1979.
Guidebook for Screening Urban Nonpoint Pollution Management Strategies. Final Report. Northern Virginia Planning District Commission, Falls Church, VA. 122 pp.
- Roffe, K.A., 1977.
Computerized Mapping for Assessment of Environmental Impacts: A Case Study. Annual Conference of the American Institute of Planners, 8-12 October, 1977, Kansas City, MO.

Appendix 1

THE
IMAGE BASED INFORMATION SYSTEM:
AN OVERVIEW

THE IMAGE-BASED INFORMATION SYSTEM:

AN OVERVIEW

The Image-Based Information System¹ (IBIS) is a computer-based approach to the analysis of geographical situations. By grouping selected IBIS programs into processing operations, a variety of spatial phenomena can be investigated. The basic concepts of data management and data processing within the Image-Based Information System are covered in this section.

IBIS is considered to be a raster-based information system

Most data entered into IBIS are in raster (image-based) format. However, the system is configured in such a manner that other data types, such as graphical and tabular data, may be used in analysis as well.

Logical and mathematical interfaces have been provided to link all data files in an IBIS data-base superstructure (Figure 2-1). By utilizing these interfaces, information may be derived from simple associations of, or comparisons between, two or more data files stored in an IBIS data base. More complex procedures including polygon overlay and cross-tabulation can also be investigated through IBIS technology.

A. DATA MANAGEMENT CONSIDERATIONS

The image-formatted data plane is the primary data type utilized in IBIS processing. IBIS data planes may be obtained directly in image form, as in Landsat imagery, or they may be derived from data compiled by sources such as the U.S. Geological Survey, the U.S. Bureau of the Census, and the Defense Mapping Agency.

Regardless of data type and origin, all data planes are incorporated into a data set that is referred to as the IBIS data base (Figure 2-2). When investigating a specific problem, a data plane may be included in, excluded from, or modified before any IBIS processing step. New data planes may be constructed with the system and may be used in subsequent processing steps.

¹The Image-Based Information System was designed by N. A. Bryant and A. L. Zobrist of the Image Processing Laboratory at JPL. The description of IBIS included in this section is quite general. A technical description of the system is included in Appendix A.

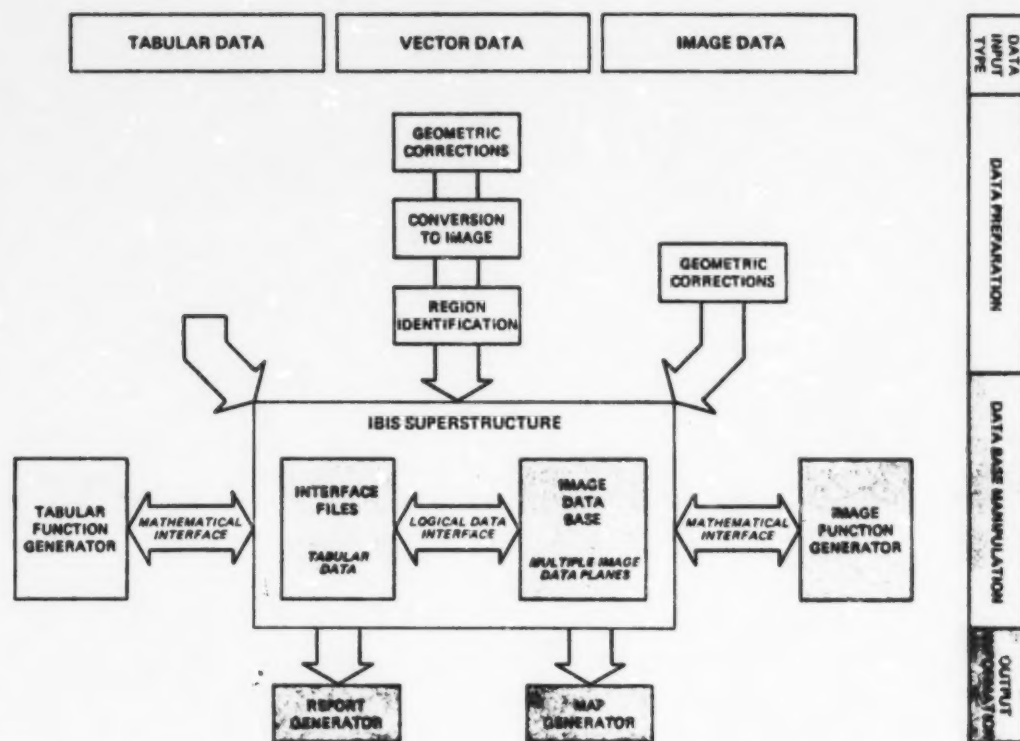


Figure 2-1. A Configuration Diagram of the Image-Based Information System

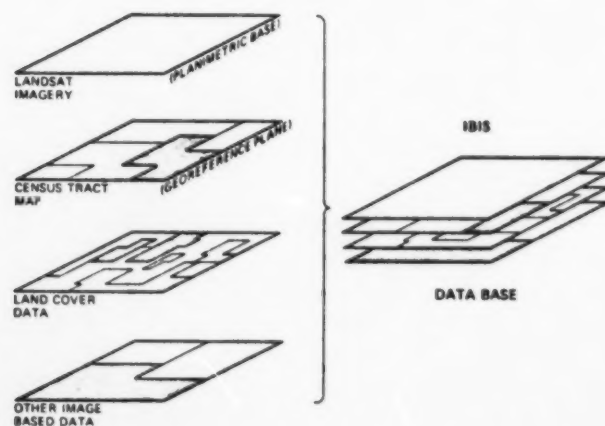


Figure 2-2. Formation of an IBIS Data Base

To maintain geometric consistency between all data planes included in the data base, an image plane exhibiting good radiometric or planimetric qualities is designated the planimetric base. Digitized topographic maps are excellent for this purpose, having been constructed from precise point datum and calibration. All other data planes are geometrically corrected to register with the planimetric base.

1. Modes of Data Input

The user of IBIS can integrate various data types to form an IBIS data base. Since the primary data structure is a raster format, image data planes are directly entered into the system. Graphical forms of data, usually obtained in Cartesian reference form, must be transformed into image space prior to inclusion as a data plane. Tabular data are not transformed into image space, but are linked to the image data base through a logical interface. Data processing requirements for each data type are unique and will be covered individually.

a. Image Data. Most image data sets entered into the data base are usually derived from Landsat imagery or other multispectral scanner sources. Other data are digitally encoded or scanned from aerial photographic products. Since image data planes are obtained from many sources, the spatial alignment of features contained in those images are often inconsistent from image to image. Consequently, provisions have been made to obtain a unified spatial surface through the implementation of resampling and spatial rectification procedures.

Geometric correction procedures and spatial rectification routines for modifying image data are features of the VICAR image processing system. Although not specifically considered to be IBIS programs, these and other VICAR programs are necessary parts of the IBIS process. VICAR software is often used to obtain special information from image data sources. Multispectral classification of Landsat imagery is one example of a VICAR procedure that produces a data set used in IBIS procedures.

b. Graphical Data. Graphical or vector data may also be entered into the IBIS data base. Graphical data are either produced locally on a coordinate digitizer, or are obtained from data tape. The Bureau of Census Urban Atlas file and DIME file are examples of data obtained on computer tapes. Regardless of data origin, graphical data are transformed into image space prior to inclusion in the IBIS data base.

As with image data, graphical data files must be in registry with the primary data base. Provisions have been made within IBIS to achieve the proper spatial alignment. (These corrections are made before the data are transformed into image space.) The registration technique utilized is a two-step process. Initially, a general surface fit is achieved through the use of a least-squares affine transformation, then an exact geometric correspondence is obtained by implementing a spatial rectification procedure. The deformation of the original surface is controlled by the selection of tie points linking geographical features that are identifiable on both the graphical data file and the primary data base.

Point and vector data are transformed into image space after geometric modifications are completed. Three-dimensional or z-value data (x, y, z) are processed in a similar manner. The Cartesian reference components of the data (x, y) are transformed into image space coordinate values, while the z-value remains unchanged.

c. Tabular Data. Tabular data may be entered into IBIS via computer cards or digital tape. These data are stored in a tabular file that is linked to the data base through a logical interface. As a result of this link, such data files are termed interface files.

2. The Georeference Plane

One of the more important graphical data files entered onto the IBIS data base is the georeference plane. The georeference plane is a polygon file used in data aggregation and map generation procedures. In one case (the Census-Urbanized Area Project), the georeference plane has been constructed from census tract data obtained from the Urban Atlas files of the U.S. Bureau of the Census.

Once the georeference plane is transformed into image space, each polygon, or region, must be identified. The region identification process involves the assignment of a unique value (or gray tone) to each individual region. After region identification, the georeference plane may be used in several higher-order IBIS procedures. For example, a polygon overlay of the georeference plane with some other image data plane can be initiated (Figure 2-3), or the gray values of each polygon in the georeference base may be modified to produce a map depicting the results of a modeling application with data stored in an interface file.

Several georeference planes may be included in an IBIS data base. For example, a data base may contain a census tract georeference plane and a congressional district georeference plane. The maximum number of regions that can be included in one georeference plane is virtually unlimited. Currently, up to 20,000 regions can be identified by gray value for an individual georeference plane.

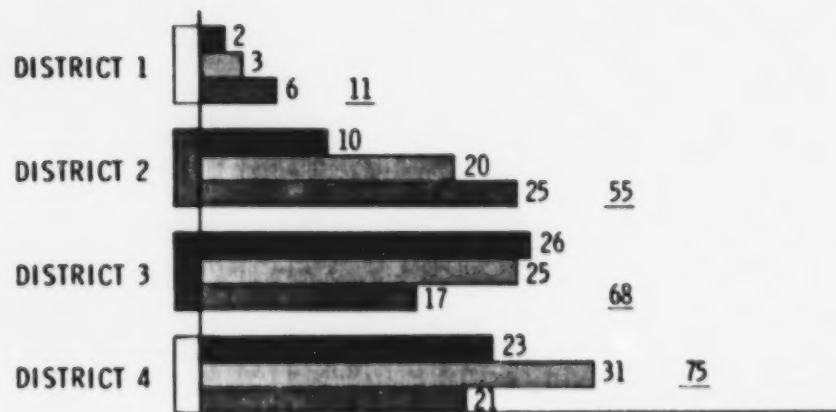
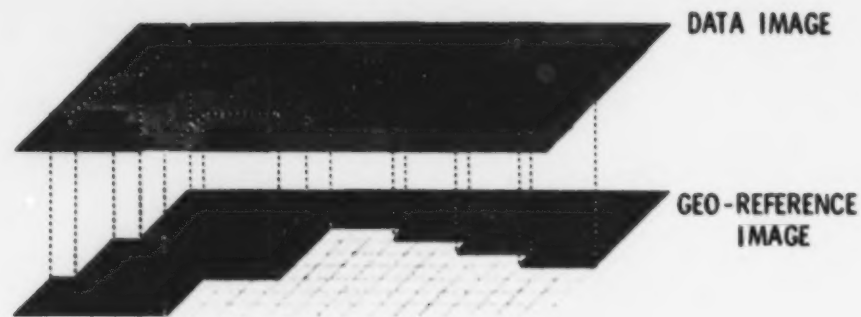
3. The Data Interface and Tabular Files

All tabular files (interface files) are linked to at least one of the georeference planes included in the IBIS data base. The specific link is obtained by storing the numerical value (gray tone) representing each region of the georeference plane with tabular data describing attributes of that region (Figure 2-4). Attribute data may be statistical in origin, an identification code, or may be the result of an image plane comparison routine such as polygon overlay or cross-tabulation.

B. MANIPULATING DATA AND OUTPUT PRODUCTS

If IBIS, or any other information system, were only a device to collect and store geographic data, the utility of the system would be quite limited. Most users of an information system require far more powerful features. Several means of data output, both pictorial and tabular, are needed. Also, the researcher may want to undertake complex modeling applications with the data files stored in the information system superstructure.

CONCEPTUAL FRAMEWORK OF IBIS INTERFACE FILE GENERATION



	TOTAL AREA	BLUE AREA	YELLOW AREA	RED AREA
DISTRICT 1	11	2	3	6
DISTRICT 2	55	10	20	25
DISTRICT 3	68	26	25	17
DISTRICT 4	75	23	31	21

Figure 2-3. The IBIS Polygon Overlay Procedure

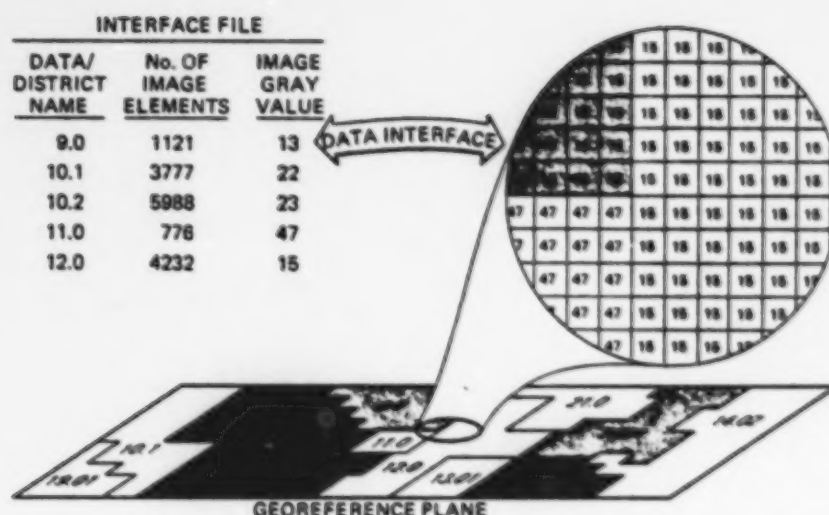


Figure 2-4. The Data Interface

Procedures for data output and data manipulation have been derived as part of the Image-Based Information System. Maps may be generated and tabular reports can be obtained.

1. Data Manipulation Procedures

Data stored in either the data base or an interface file can be modified or manipulated with IBIS software. New data planes and interface files are easily generated. Four basic data-manipulation procedures are currently available.

a. Data Manipulation Between Image Planes. New image data planes are generated as a function of two or more image data planes. Chiefly, the procedures implemented to derive such data planes are VICAR routines, although some IBIS routines are also used.

Simple transformations such as image addition, subtraction, multiplication, and division are easily obtained. Complex functions are handled nearly as easily, and precise mathematical formulas may be specified. Image enhancement routines are available, as are several data classification and stratification routines.

b. Data Manipulation Within the Interface File. Most functions available in the image domain are also available for analysis of tabular data. Resultant from such operations, new tabular data entries are generated. Complex mathematical functions may be used to derive higher-order properties of data stored in an interface file.

c. Data Manipulation of Image Data into Tabular Data. By implementing certain IBIS routines, data originally stored in image format may be summarized and copied into a tabular file. The majority of these routines are aggregation functions, an example of which is histogramming.

IBIS procedures for polygon overlay and cross-tabulation are within this realm of data transfer programs.

d. Data Manipulation of Tabular Data into Image Data. The representation of tabular data in image form is primarily used as an output aid. By the implementation of a map generating routine, any georeference base can be modified as a function of an interface file. Modeling of data is performed similarly. Data planes produced in this manner can be entered into the IBIS data base for subsequent operations.

2. Data Output Features

Two output formats are available to the system user: maps and tabular reports. Maps are produced directly from any image data plane or through modification of a georeference plane. Tabular reports are made available through the operation of a report generator.

C. SUMMARY

With a knowledge of image processing, an analyst can learn to operate the IBIS system. A researcher can utilize the system to store several data planes and much tabular data. With all of the information at the data user's disposal, many complex modeling problems may be solved relatively effortlessly.

The various modes of data entry, data manipulation, and data output provide the researcher with complete flexibility to structure a unique data base specifically designed for a particular problem or investigation. IBIS is merely a framework for analysis of spatial data. The actual information system is constructed with the selection of specific image and tabular data.

Appendix 2

LANDSAT AND IBIS
PROCESSING STEPS
FOR THE
OCCOQUAN DATA BASE

Landsat and IBIS Processing Steps for the Occoquan Data Base

Overview

NASA Kennedy Space Center (KSC) in an interagency agreement with the Environmental Protection Agency (EPA) in Athens, Georgia, assumed management of a remotely sensed data base for the HSPF modeling of the Occoquan Basin. In addition, KSC, with in-house personnel, provided the land use classifications for this data base through the Image 100 interactive computer image processing techniques on Landsat digital data. Under contract to KSC, the Jet Propulsion Laboratory (JPL) used their Image Based Information System (IBIS) to relate a Northern Virginia Planning District Commission (NVPDC) furnished ground sensed data base to the KSC Landsat analysis. This ground sensed information on slope, permeability, and soil characteristics coupled with the Landsat classifications created a novel and cost effective data base for the assignment of parameters to the HSPF water quality model.

Landsat Data Selection and Preparation

Hydrocomp personnel concluded that 1975 was a good "event" year for running the HSPF with remote sensing data because during this year the Occoquan Basin received about normal rainfall and water quality and stream flow data had been collected which could validate model predictions. It was also concluded by the NVPDC staff and a Prince William County Agriculture Extension Agent that the best dates for separation of ground cover are September 15, April 15, and July 1, in order of degradation of separation.

Four Landsat tapes with minimum cloud cover were obtained in these time periods. These four sets of tapes were dated March 12, 1975, April 8, 1975, August 3, 1975, and October 14, 1975. These tapes were geometrically corrected and a final tape for each date was made having a pixel size of 79 x 79 meters and a total area of 910 x 910 pixels. Because the Image 100 CRT has a format of 512 x 512 pixels, each of the corrected tapes were separated into four sections, NE, NW, SE, and SW, of 512 x 512 pixels.

A velum map tracing of the Occoquan sub-basins was made from a 1:48,000 scale NVPDC land cover map. This velum map was photo reduced to 1/2 size and the various sub-basins colored red, green, and blue. This sub-basin map was then overlaid onto the previously developed Image 100 Landsat scenes using the TV camera input and the red, green, and blue filters. These sub-basins were placed into the Image 100 Channel 5 with each basin being assigned to one theme channel. Now each corrected Landsat scene had a sub-basin overlay.

Image 100 Analysis

The initial step was to determine the best temporal combination of various dates to achieve maximum separability of classes. The March 1975 Landsat data was excluded because it did not indicate a significant probability of separation of the various ground cover classes. The April, August, and October Landsat data were used to obtain radiance values for various classes. Maximum radiance ranges for each Landsat band were evaluated and it was concluded that the optimum temporal combination was to combine Landsat bands 6 and 7 from the April 8, 1975, data and Landsat bands 4 and 5 from the August 3, 1975, data. This combination was the data used for this project.

In a field trip to the Occoquan Basin by KSC personnel some twenty-three ground truth classes were established in over 400 ground truth points. Signatures were developed interactively on the Image 100 for each of the ground truth classes. Due to the heterogenous nature of the Occoquan Basin, there were overlaps of classes in many cases and a set of ten multi-nodal signatures were developed which did not have any noticeable overlap. To meet Image 100 format, these 10 signatures were merged into a final set of eight signatures. Each one of the final eight signatures was fed into one of the theme tracks of Channel 5 with all other seven theme tracks containing a sub-basin overlay. Then a function was performed between the theme and each sub-basin which erased the theme in all areas except where it overlapped the specific sub-basin. A printout of the number of pixels of that specific classification was made for each of the sub-basins. Finally, a calculation was made to convert the number of pixels to area in acres.

IBIS Data Preparation

The KSC thematically classified data with sub-basin overlay and the Occoquan Basin ancillary data compiled by NVPDC was forwarded to JPL for reformatting to an IBIS data base. NVPDC, through ground surveys and data from other agencies, compiled slope, soils, permeability, and general land use data on the Occoquan Basin into the GRIDS system developed by Harvard University. This ancillary data was averaged over squares of 1,000 x 1,000 feet. The initial step by JPL was to reconstruct the KSC and NVPDC data into a Video Image Communication and Retrieval (VICAR) system compatible format. After the ancillary data was converted to image-based format, a variety of images were produced of the Occoquan Basin showing soil types, slope magnitudes, and the KSC polygon sub-basin overlay and land classifications.

Landsat planimetric controls were established for the registration of GRIDS to the Landsat data which was averaged over picture elements of 79 x 79 meters. Upon receipt of a subsequently supplied Occoquan Basin base map with grid cell overlay, JPL digitized a set of longitude/latitude control points. Registration of the GRIDS data to the Landsat data was then accomplished. The final preparation step was the development of the 15 sub-basin image-based map to be used as regions of statistical aggregation. Although only sub-basins 4, 5, and 7 were considered in this study, JPL produced thematically classified Landsat acreages and GRID data cell tabulations for all 15 sub-basins.

Final IBIS Data Base Products

The IBIS data base was queried in numerous cross correlations and the following information was obtained.

1. Acreage tabulation and pixel count by 100 soil types for each of the eight Landsat land use classifications.
2. Acreage tabulation and mean values for each land use for the following:
 - a. Slope for seven percent intervals
 - b. Volume of available storage in eight categories
 - c. Restrictive rates in one-tenth intervals
 - d. Erodibility "K" in one-tenth intervals
3. Acreage tabulation and percentage distribution for each land use for the following:
 - a. Hydrologic soil groups for four categories
 - b. Soil texture for six categories
 - c. Volume of available storage for eight categories
 - d. Restrictive rate for one-tenth intervals
 - e. Erodibility "K" for one-tenth intervals

4. Acreage tabulations for volume of total water above the restrictive layer in one-tenth intervals for each land use.